2019

THE FUNCTIONAL ROLE OF RNA BINDING PROTEIN RBMS3 AS A TUMOR PROMOTER IN TRIPLE-NEGATIVE BREAST CANCER CELLS

Yuting Zhou
University of Kentucky, yuting.zhou@uky.edu
Author ORCID Identifier:
https://orcid.org/0000-0002-6768-9338
Digital Object Identifier: https://doi.org/10.13023/etd.2019.433

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Yuting Zhou, Student

Dr. Binhua P. Zhou, Major Professor

Dr. Trevor Creamer, Director of Graduate Studies

The Functional Role of RNA Binding Protein RBMS3 as a Tumor Promoter in Triple-Negative Breast Cancer Cells

---

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Medicine
at the University of Kentucky

By
Yuting Zhou

Lexington, Kentucky

Director: Dr. Binhua P. Zhou, Professor of Molecular and Cellular Biochemistry

Lexington, Kentucky

2019

https://orcid.org/0000-0002-6768-9338

ABSTRACT OF DISSERTATION

THE FUNCTIONAL ROLE OF RNA BINDING PROTEIN RBMS3 AS A TUMOR PROMOTER IN TRIPLE-NEGATIVE BREAST CANCER CELLS

RBMS3 belongs to the family of *c-myc* gene single-strand binding proteins (MSSPs) that play important roles in transcriptional regulation. Here, we show that RBMS3 functions as a tumor promoter in triple-negative breast cancer (TNBC), a highly aggressive BC subtype. Analysis of RBMS3 expression shows that RBMS3 is upregulated at both mRNA and protein levels in TNBC cells. Functionally, overexpression of RBMS3 increases cell migration, invasion and cancer stem cell (CSC) behaviors. Moreover, RBMS3 induces expression of epithelial-mesenchymal transition (EMT) and CSC markers. Conversely, loss of RBMS3 in TNBC BT549 cells inhibits cell proliferation, migration and mesenchymal phenotype. Correlation analysis shows RBMS3 is associated with TGF-β signaling. Mechanistically, RBMS3 interacts with *Smad2*, *Smad3* and *Smad4* mRNA and regulates the stability of these transcripts. Importantly, RBMS3 prevents TGF-β-induced cytostasis and apoptosis in premalignant cancer cells. Moreover, RBMS3 inversely correlates with expression of ESRPs, epithelial-specific splicing regulatory proteins that regulate morphogenesis-associated alternative splicing events. ESRPs suppress EMT through distinct mechanisms: ESRP1 restricted cell migration, whereas ESRP2 prevented cell growth. RBMS3 significantly facilitates the EMT process when ESRPs are lost. Collectively, the studies within this dissertation identify RBMS3 as a positive regulator of EMT and breast cancer progression by regulating the TGF-β signaling pathway.

KEYWORDS: TNBC, RBMS3, EMT, TGF-β signaling, mRNA stability, and ESRP

Yuting Zhou

11/09/2019

THE FUNCTIONAL ROLE OF RNA BINDING PROTEIN RBMS3 AS A TUMOR PROMOTER IN TRIPLE-NEGATIVE BREAST CANCER CELLS

By
Yuting Zhou

Binhua P. Zhou, Ph.D.
Director of Dissertation

Trevor Creamer, Ph.D.
Director of Graduate Studies

11/09/2019

Dedication to my family, wholeheartedly.

ACKNOWLEDGMENTS

First and foremost, I must acknowledge and thank my Dissertation Chair, Dr. Binhua P. Zhou, an extraordinary scientist, wise advisor and kind friend. He provided invaluable guidance, instructive comments and boundless support at every stage of the dissertation process. He also encouraged me to expand my knowledge in the cancer research field and granted me the freedom to test out my own ideas as part of the training process. Undoubtedly, Dr. Zhou has been the most influential mentor in my life and I could not be more thankful that I was able to work and study in his laboratory for the past five years. As I move on to the next chapter of my career, know that he has my sincerest gratitude for pushing me to be the best I can be. And wherever I go, I will always feel humbled and proud to say that Dr. Peter Zhou is my Ph.D. mentor.

Next, I would like to express my heartfelt gratitude to each member of my Dissertation Committee: Dr. Tianyan Gao, Dr. Yvonne Mittendorf and Dr. Qing-Bai She. I could not be more honored to have Dr. Gao's encouragement and advice to keep me on track to obtain my Ph.D. I earnestly admire her intelligence, capability, strength and charisma. I also owe a big thanks to Dr. Mittendorf, who kindly provided guidance on how to blend into the research community as an international student. I feel equally grateful to Dr. She for taking time from his busy schedule to critique my proposal writings and gave me instructive feedback. I truly appreciate that they have offered insights that guided and challenged my thinking in order to shape my dissertation project. A warm thank you goes to my outside examiner: Dr. Christopher Richards.

I would like to say thanks to all present and past members of Dr. Zhou's laboratory. I am much obliged to Dr. Zhibing Duan, Dr. Weijie Guo, Dr. Jian Shi and Dr.

Yiwei Lin. Each have lent their expertise and years of help to push my project forward. I honestly feel grateful to have such incredible lab mates and friends, who are always willing to help and support each other. In particular, I deeply appreciate Dr. Duan for supporting and instilling in me the spirit of determination, strength, devotion and sheer tenacity. Thanks also goes out to Dr. Yadi Wu for her words of advice on my career path.

For the past few years, I have received continual support from members of the Markey Cancer Center and Department of Molecular and Cellular Biochemistry. Many thanks to our Director of Graduate Studies, Dr. Trevor Creamer, who guided me through qualify process and made every endeavor to make sure I stay on track. I also wish to extend special thanks to Dr. Kathleen O'Connor for sharing her love and passion for life and science as well as for her time and effort to critique my proposal.

Sincere appreciation for the beautiful people and friends I met at the University of Kentucky: Dr. Hong Pu, Fei Xiong, Lu Dai and Jianlin Wang. I thank them for the good times we have shared together. My best friends, Hanxuan Huang and Jiaxin Li, have my immense thanks for all the support they have given me during my worst times. Forever in my heart is my soulmate Shuyi Li, who will always be blessed and missed.

Finally, words are powerless to express how indebted I feel towards my dear parents and my brother Kang. They are always there for me when I encounter hardship and frustration. I could not have come this far without their endless love and support. Every once in a while I see myself being with them in my sweet dreams, heaven knows how much I miss them! And I must also express my love and gratitude to the love of my life, Raku and Albert. Lots and lots of love!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

1.1 Mammary Epithelium

About 80% of all breast cancer (BC) are invasive ductal carcinoma. Mammary epithelial is primarily comprised of two cell types: luminal and basal. The luminal/epithelial cells constitute the inner layer of the ducts and lobuloalveolar units, whereas the basal/myoepithelial cells form the outer basal layer. The epithelial duct is enveloped by a layer of basement membrane (BM), a specialized form of extracellular matrix rich in collagen IV and laminin. BM contains a repertoire of membrane-tethered growth factors that can be released by tumor cell-secreted proteases. BM also participate in signal transduction events via integrin-mediated cell-matrix adhesion, leading to alterations in cell polarity, invasiveness, proliferation and survival. It is embedded within a complex stroma, which contains fibroblasts, adipocytes, blood vessels, nerves and various immune cells [1]. It is well documented that luminal epithelial cells establish apical-basal polarity during differentiation process. While the apical side is exposed to the lumen, the lateral surface is intimately associated with adjacent cells and basal cells through cell-cell contact structures, i.e. tight junctions, adherens junctions and desmosomes This organization is crucial for the structural integrity of the epithelia [2].

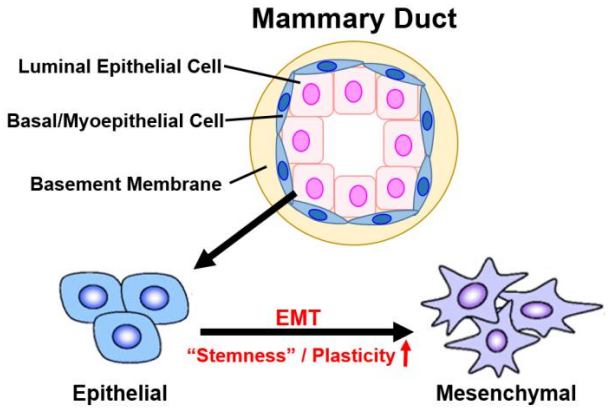


Figure 1.1 Schematic on the main components of mammary duct
Activation of the EMT program in non-metastatic epithelial breast tumors confers mesenchymal phenotypes and cancer stem cell traits to cancer cells.

1.2 Breast Cancer

Statistics

Breast cancer (BC) is the leading cause of cancer-related deaths among women [3]. Approximately 268670 new cases of invasive BC and 41,400 deaths are projected to occur for the year 2018; this mortality is primarily due to metastasis to other organs, preferentially the lungs and bones [4, 5]. According to American Cancer Society, there will be an estimate of 25,990 new cases and 10,590 deaths in Kentucky during the year of 2018 [6]. Although the number of breast cancer cases has increased over time, breast cancer incidence rates have been fairly stable for the past 10 years. Early diagnosis of the disease can lead to a good prognosis and a high 5-year survival rate of up to 90% [7].

Risk Factors

Numerous risk factors are associated with breast cancer, including gender, aging, family history, gene mutations, ethnicity, estrogen receptor (ER) status and unhealthy lifestyle [7]. Unsurprisingly, breast cancer cases are 100 times more prevalent in women than in men. While most breast cancers are found in women who are 50 years old or older, about 11% of all new cases of breast cancer in the United States occur in women younger than 45 years of age. About 15% of breast cancer patients have family history. The risk of getting breast cancer nearly doubles if a woman has a first-degree relative (mother, sister or daughter) who has been diagnosed with breast cancer. 5-10% of breast cancers can be linked to gene mutations, such as *BRCA1* and *BRCA2* gene mutations, inherited from one's mother or father. For instance, women with a *BRCA1* mutation have a 55-65% lifetime risk of developing breast cancer. About 85% of breast cancer occur in women without family history, due to genetic mutations that happen as a result of the aging process and lifestyle in general, rather than inherited mutations. Female breast cancer incidence and

mortality rates vary significantly by ethnicity in the United States. Non-Hispanic white (NHW) and Non-Hispanic black (NHB) women have higher incidence and death rates than women of other ethnicities [6]. While NHW women develop breast cancer slightly more than NHB women, however, NHB women tend to die of breast cancer more often, possibly due to the preference for developing a more aggressive type of tumor. NHW women have the highest rates of ER^+ breast cancer, whereas NHB women have highest rates of ER^- breast cancer [8]. According to Johns Hopkins Medicine Health Library, lifestyle-related risk factors for breast cancer include physical inactivity, poor diet, obesity, frequent alcohol consumption, breast irradiation, no or late pregnancy (after age 30), early menstrual periods (before age 12), late menopause (after age 55), recent use of oral contraceptives, long-term use of combined hormone replacement therapy and exposure to pesticides or other chemicals.

Molecular Classification

Based on molecular classifications, breast cancer can be divided into at least five intrinsic subtypes: HER2-enriched, luminal A, luminal B, basal-like and normal breast-like [9]. HER2-enrich subtype is characterized by overexpression of the HER2 oncogene and other genes pertaining to the HER2 amplicon. Luminal class comprises the majority of breast cancer and is further stratified into two subtypes. Luminal A subtype is defined as ER+, PR ≥20%, HER2-, Ki67 <14%, and, if available, 'low' recurrence risk. Luminal B subtype is ER+, HER2-, and at least one of the following: Ki67 ≥20%, PR- or <20%, and, if available, 'high' recurrence risk [10]. Luminal B tumors are associated with worse prognosis than that of the luminal A subtype. Generally, the luminal class is characterized by high proliferation rates, with ER expression, and/or HER expression and low /absent PR expression. The basal-like subtype is largely defined by lack of expression of ER and HER2, positive expression of genes characteristic of basal-like cells of the breast and high

proliferation index. Normal breast-like subtype displays a triple-negative phenotype (ER^-, $HER2^-$ and PR^-) and an expression profile similar to those found in normal breast tissue [11].

1.3 Basal-like Breast Cancer vs. Triple-Negative Breast Cancer

Basal-like breast cancer (BLBC) is a highly malignant subtype of breast cancer. A large subset of this subtype is characterized by lost expression of ER, HER2 and PR and high expression of several basal markers, including cytokeratins (5/6, 14 and 17), EGFR, caveolin 1 and P-cadherin, exhibiting a 'triple-negative' phenotype [12]. Triple-negativity renders this subset of patients less responsive to hormone therapies. Undoubtedly, tremendous scientific advancements have been made to increase survival rates of BC patients, however, this is only applicable to BC diagnosed early stages without metastasis.

BLBC accounts for up to 20% of all human breast carcinoma. While approximately 70% of BLBC are TNBC, about 75-80% of TNBC are BLBC. BLBC with 'triple-negative' phenotype (TNBC) tend to be aggressive and destructive as it is associated with a high proliferation index, ensuing brain and lung metastasis, and poor clinical outcome in spite of treatment [13]. Compared to other molecular subtypes, BLBC still carry the worst prognosis due to decreased disease-free survival, disease-specific survival and overall survival. In addition, BLBC patients are at increased risk for early relapse or recurrence within 3-5 years of diagnosis. Owing to its aggressive phenotype and poor prognosis as well as limited response to endocrine therapy, effective treatment for BLBC has been an extremely difficult challenge for breast cancer researchers and clinicians worldwide [14].

1.4 Therapeutics

Traditionally, aggressive breast cancers have been treated with general chemotherapy drugs anthracycline and paclitaxel. While being more sensitive to this

regimen in comparison to luminal and normal-like subtypes, over 50% of the BLBC patients continues to have residual disease and carry a high risk of relapse within 5 years of diagnosis. Besides, general chemotherapy is limited by nonspecific cytotoxicity, which poses significant health concerns to patients [14]. Unlike other subtypes, BLBC lacks expression of molecular targets that responsiveness to effective hormone therapies [15]. Hence, research and clinical trials for targeted therapies are being carried out in the hope of achieving better therapeutic indices.

DNA repair pathway: BRCA1 mutation has been identified in 15-20% of women with a family history of breast cancer[16]. BRCA1 deficiency is commonly observed in BLBC, which provides great opportunity to targeted therapy. BRCA is responsible for homologous recombination repair of double-strand breaks. Platinum-based chemotherapy (including cisplatin and carboplatin) induces DNA cross-linking, resulting in double-strand breaks in cells. Cancer cells with BRCA deficiency become vulnerable to apoptosis upon treatment of platinum agents. Since poly(ADP-ribose) polymerase (PARP) is involved in base excision repair for single-strand breaks, PARP inhibitors can be combined with platinum agents to create more DNA lesions that are less likely to be repaired by the DNA repair machinery. This cooperative approach utilizes the concept of 'synthetic lethality' by targeting complementary pathways leading to a lethal combination [14]. In 2018, 5 PARP inhibitors (talazoparib, nirapari, rucaparib, olaparib and veliparib) have finally become available for patients with BRCA-mutant metastatic breast cancer.

Aberrant activation of the VEGF pathway shown in BLBC has led to therapeutic strategies to target VEGF and its receptors, including anti-VEGF antibody (Bevacizumab), VEGFR inhibitor and other receptor tyrosine kinase inhibitors (Sunitinib, etc.).

Although EGFR is upregulated in most BLBCs, the use of a dual inhibitor of EGFR and HER2/neu (Lapatinib) does not seem to benefit $HER2^-$ BLBC patients. Inhibitors of other downstream kinases, such as MEK and PI3K, are also considered for targeted therapies[14].

Within the past ten years, considerable evidence has highlighted the importance of immune response in influencing the progression of TNBC. Presence of tumor-infiltrating lymphocytes (TILs) has been widely recognized as a prognostic factor in early-stage TNBC. Also, expression of immune evasion molecules in the tumor microenvironment, such as programmed death-ligand 1 (PD-L1), may influence TNBC prognosis. In TNBC, PD-L1 expression is about 40-65% in immune cells. Most patients tested as PD-$L1^{+}$ in immune cell tumors also had positive PD-L1 expression on tumor cells [17]. Development of new therapeutic agents against immune checkpoint molecules, such as anti-PD-1 and anti-PD-L1 monoclonal antibodies has emerged, bringing BC into immunotherapy era [18].

1.5 Breast Tumor Progression and Metastasis

Like many other tumors, breast tumors progress from an early pre-neoplastic lesion to the development of clinically detectable distant metastatic foci, which involves a series of genetic and epigenetic alterations affecting both tumor cells and the surrounding stroma. Just like other evolutionary processes in nature, the path toward metastatic colonization is not only extremely complex, but also highly inefficient. Most cells that leave a tumor fail to seed distant organs and often die on the path [19]. To develop metastasis, epithelial cells in primary tumor must invade locally and escape from the physical barriers (extracellular matrix and basement membrane), intravasate into the circulation (lymphatic or vascular system), extravasate into the parenchyma of distant tissues, survive in the foreign milieu and re-initiate proliferation programs to form overt metastases [20, 21]. The metastatic potential of individual cells within the bulk of a tumor is thought to be widely heterogeneous, which is influenced by various factors, such as tumor size, cell of origin and type of oncogenic driver mutation [22].

1.6 Epithelial -Mesenchymal Transition

Overview

During development and morphogenetic events, epithelial cells undergo a process called epithelial to mesenchymal transition, or EMT, the concept of which was first described in the 1980s [23]. During EMT, cells lose epithelial characteristics, such as intrinsic polarity, cell-cell contact and cell-matrix contact, but gain mesenchymal properties, including fibroblastoid morphology, increased proliferation and motility, allowing them to break through the basement membrane and invade surrounding tissues, even to distant organs [24]. Invasion is a critical step to progression toward a malignant disease. Morphologically, during the initiation of EMT basal-membrane anchored cells transform from a cuboidal epithelial-like cell to a spindle-shaped mesenchymal-like cells [25]. Meanwhile, adherens junction complexes dissemble and the actin cytoskeleton reorganize from an epithelial cortical alignment into actin stress fibers, which are anchored to focal adhesion complexes. E-cadherin, a major component of the adherens junctions, binds tightly to β-catenin with its cytoplasmic domain, which in turn anchors to the actin cytoskeleton via acting-binding proteins. Loss of E-cadherin epithelial marker from the basolateral membrane is associated with release of membrane-bound β-catenin to the cytosol and activation of the canonical Wnt pathway to regulate pluripotency factors Oct3/4, Sox2 and Nanog, which promotes cell growth, survival and maintenance of stemness [26]. Loss of E-cadherin, accompanied by induction of mesenchymal markers (Snail, Slug, Twist1, Zeb1/2, N-cadherin and Vimentin, etc.), represent hallmarks of EMT and tumor progression. These phenotypic traits appear to be the fundamental molecular basis for tumor initiation and expansion, metastatic capacity, tumor recurrence and therapy resistance [27].

The EMT program is activated through contextual signals that cells receive from their neighbors. In the case of carcinoma, these signals derive primarily from the

fibroblasts, mesenchymal stem cells and inflammatory cells that are recruited to the stroma of the tumors and contribute to the formation of a tumor-promoting microenvironment [28].

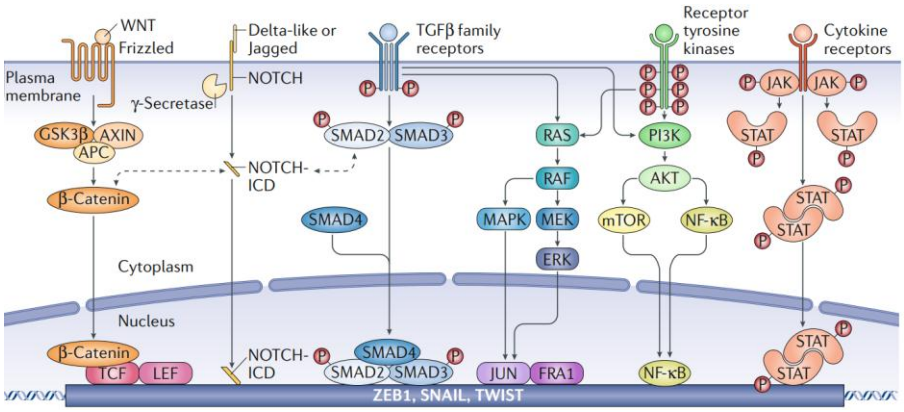


Figure 1.2 Core signaling pathways that activate EMT

In the context of non-neoplastic cells, several cell-intrinsic signaling pathways (TGF-β, WNT and NOTCH, etc.) are activated, upon binding of specific ligands, cytokine or growth factors to their cognate receptors, to induce EMT during embryonic development and wound healing. Canonical WNT pathway is activated upon binding of WNT ligands to the Frizzled family of membrane receptors, which leads to the release of β-catenin from the GSK3β-AXIN-APC complex. β-catenin enters the nucleus and binds to transcription factors (TF) TCF and LEF to activate genes that drive EMT. The NOTCH pathway is stimulated upon binding of the Delta-like or Jagged family of ligands to the NOTCH receptor, which triggers proteolytic cleavage and release of the active, intracellular domain of the NOTCH receptor (NOTCH-ICD), which translocates into the nucleus to function as a transcriptional co-activator. TGF-β pathway also crosstalk with PI3K-Akt pathway to trigger the activation of mTOR complex and NF-κB, p38 MAPK pathway and the RAS-RAF-MEK-ERK signaling axis. Several cytokines may trigger the phosphorylation, dimerization and activation of JAK and signal transducer and activator of transcriptional proteins, which also activate transcription of EMT transcription factors [29]. Details of TGF-β pathways is discussed below. (Figure cited from Dongre et al.[29])

Classification

EMT is classified into three types based on distinct biological settings under which they occur, leading to various functional consequences. Type 1 EMT occurs during implantation, embryo formation and organ development to generate diverse cell types that share common mesenchymal phenotypes. This class of EMT does not cause fibrosis or induces an invasive phenotype resulting in systemic spread of the cell through the

circulation. Type 2 EMT responds to wound healing, tissue regeneration and organ fibrosis. Following injury and inflammation, this program is activated as a form of repair event that gives rise to fibroblasts and other related cells in order to reconstruct the tissues [30]. However, persistent EMT caused by chronic inflammation may ultimately lead to organ destruction. Type 3 EMT is found in neoplastic epithelial cells at the invasive front of primary tumors that undergo different extents of phenotypic conversion to invade and metastasize through the circulation and generate life-threatening metastatic lesions at distant tissues and organs. While these EMT events are involved in considerably different biological processes, some genetic elements and regulatory mechanisms may be similar or well-conserved [24].

Single Cell Migration

Single cells utilize two major modes of migration: amoeboid and mesenchymal. Amoeboid migration features blebbing, weak adhesions and rapid motility, whereas mesenchymal migration exhibits extensive stress fibers, polarization and a leader-trailer edge [31, 32]. Amoeboid migration primarily occurs when cells deform to pass through pores and fibers in the ECM, whereas mesenchymal migration is accompanied by creation and expansion of paths as cells degrade and remodel the surrounding matrix [33]. Besides using matrix metalloproteases (MMPs), cells also rely on Rho kinase, integrin and actomyosin to enable entry into pores via deformations to the cytoskeleton [34]. The mode of migration employed by a cell is partially mediated by its adhesivity to the matrix and the architecture, porosity, composition and mechanical properties of the ECM. For example, intrinsic activity of actomyosin is the key to transmitting mechanical signals from the ECM to the cell and generation of contractile force within the cytoskeleton that is transmitted to adhesion complexes to facilitate movement of cells along the matrix. While these adhesion complexes are essential for mesenchymal migration, they are less important for amoeboid migration [31, 35].

Collective Migration

EMT has traditionally been described as a binary process that involves complete conversion from epithelial to mesenchymal state. However, recent work points to a greater flexibility in this transitional process, as increasing evidence reveals that EMT may subsume a spectrum of intermediate or 'hybrid' states, which has been commonly referred to 'partial EMT' [36]. For example, migratory neural crest cells, particularly those arise in the Xenopus and fish embryos, migrate in a collective manner but have not undergone a full EMT. These cells gain migratory and invasive properties but have not lost epithelial properties, as they maintain a significant degree of cell-cell adhesion and cadherin expression that allow them to migrate together in a coordinated manner using a variety of modes including sheets, strands, tubes and clusters [32]. Similarly, metastatic carcinoma cells from spontaneously-arising tumors could have activated EMT program but never become entirely mesenchymal by completing this process [37].

Overall, two main types of mechanisms exist in cancer cell invasion and dissemination: the first type may give rise to single cells capable of crossing basement membrane and invading circulation; the second type where cells migrate in a multi-cellular cluster by retaining cell-cell contact. While both situations have been observed in mouse models and clinical specimens, the mechanism underlying these differences in epithelial plasticity remains elusive. It appears that divergent EMT programs correlate with tumor subtype in several human breast carcinoma lines. Typically, in the well-differentiated luminal A, luminal B or normal-like BC cell lines, EMT is associated with persistent expression of *CDH1* mRNA and re-localization of E-cadherin protein, which features a partial EMT program [36]. Conversely, cell lines harboring a less differentiated basal signature and loss of *CDH1* mRNA and E-cadherin protein features a complete EMT program (Figure 2). Aiello et al found that partial EMT cell lines are predisposed to form circulating tumor cell (CTC) clusters more readily than complete EMT cell lines, consistent

with findings of Giampieri et al., indicating that TGF-β signaling in BC cells prompted a switch from collective migration to single-cell migration, potentially attributable to conversion from partial to complete EMT program [38]. CTC clusters showed enhanced metastatic potential compared to single cells, Aiello et al. speculated that tumors exhibiting a partial EMT phenotype might exhibit an increased metastatic rate than tumors exhibiting a complete EMT phenotype[36, 39-41]. Likewise, tumor cells with various degrees of epithelial-plasticity are prone to metastasize to different sites. Due to complexity and multimodality in the factors governing metastasis, the relationship between partial EMT and complete EMT and their respective clinical outcome remain to be defined [42].

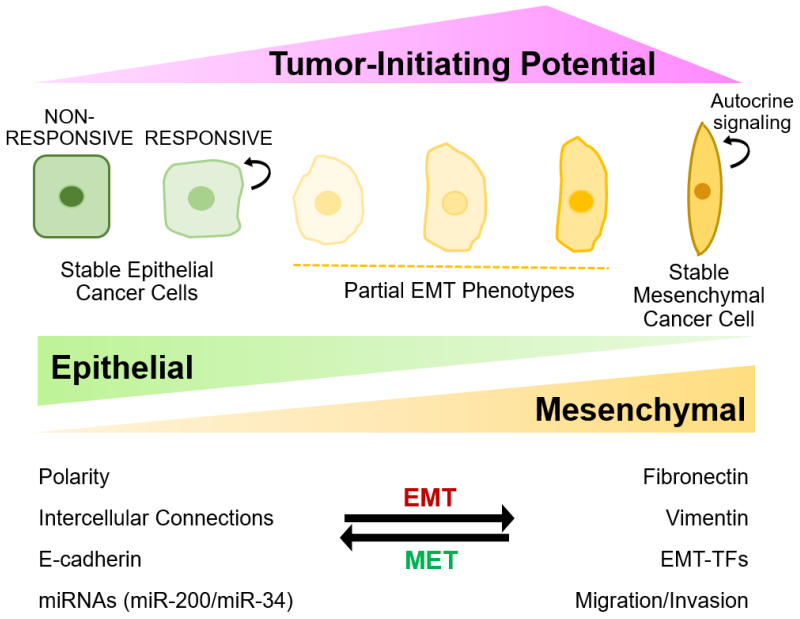


Figure 1.3 An epithelial-mesenchymal spectrum exist in breast cancer

Epithelial breast cancer cells that fail to respond to EMT-inducing signals are unable to undergo EMT, whereas responsive epithelial cancer cells exhibit disrupted autocrine signaling and transition toward mesenchymal cancer cell state. Once the transition is complete, these cells may maintain a stable mesenchymal phenotype by autocrine signaling in the absence of the EMT-inducing signals [43]. The EMT process is thought to be reversible and is associated with tumor-initiating potential, which peaks at partial-EMT

phase and diminishes as cells reach a stable mesenchymal state. Listed are some of the factors that preferentially reside in either cell state [44]. (EMT-TF, EMT-transcription factors; Figure modified from Chaffer et al. [45])

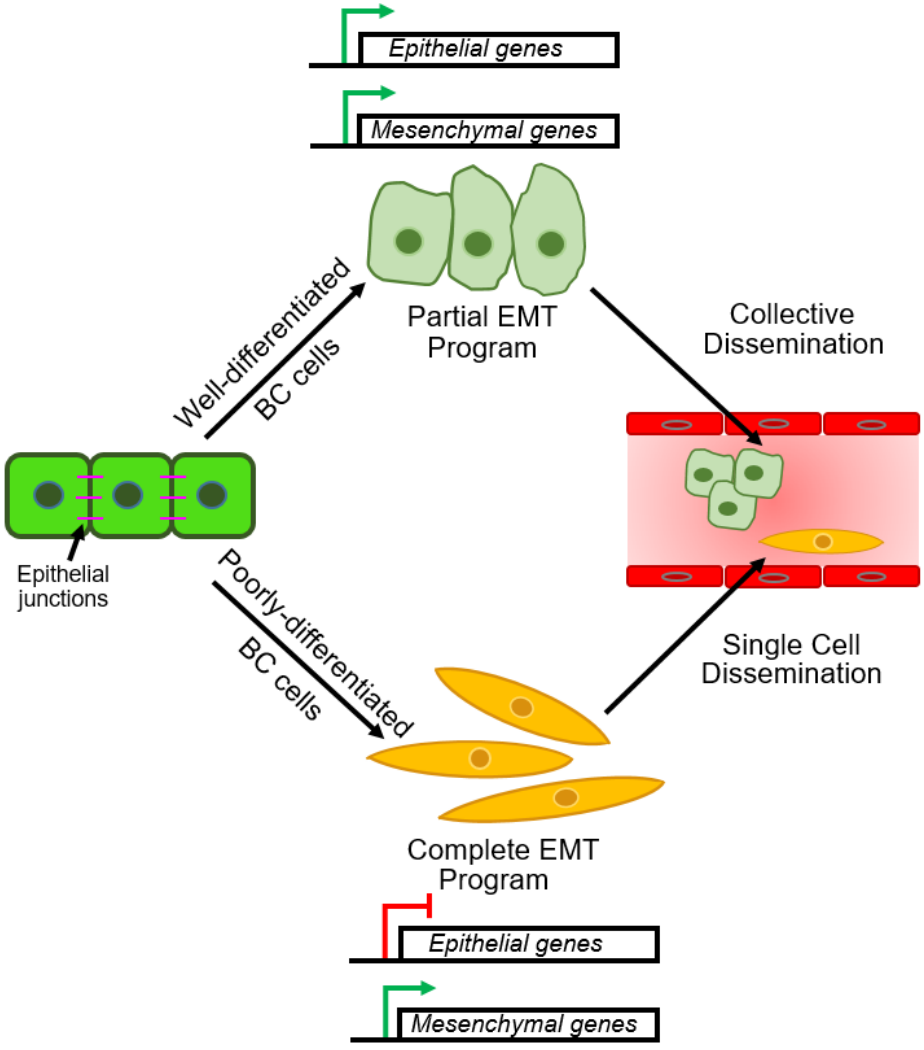


Figure 1.4 Schematic of differences between partial and complete EMT programs
This figure models two types of tumor cell dissemination. Poorly-differentiated BC cells are prone to activate a complete EMT program by transcriptional repression of epithelial genes and activation of mesenchymal genes. Conversely, well-differentiated BC cells, as observed in most epithelial tumors, modify their epithelial phenotype through an alternative program involving protein internalization rather than transcription repression, leading to a 'partial EMT' phenotype. As a result, cancer cells utilize this mechanism to migrate as clusters in the circulation, contrasting with the single-cell migration pattern defined by traditional concept of 'EMT'. (Figure modified from Aiello et al. [36])

Transcriptional Regulation

Most signaling pathways leading to EMT induction converge on the down-regulation of E-cadherin, a critical epithelial cell adhesion molecule that serves as the gatekeeper of EMT [46, 47]. In most cell types, loss of functional E-cadherin results in decreased cell adhesion, leading to uncontrolled cell growth and metastasis.

Molecular events of EMT are transcriptionally controlled by master regulators such as transcription factors Snail, Slug, ZEB1/2 and Twist1/2. In most physiological EMT conditions, overexpression of Snail, Slug, ZEB1/2, or Twist1/2 in epithelial cell lines typically induces EMT [25]. Snail and Slug initiate EMT by repressing epithelial genes like E-cadherin through binding to E-box DNA sequences of their carboxy-terminal zinc-finger domain, while ZEB1/2-mediated transcriptional repression often requires additional recruitment of a C-terminal-binding protein (CTBP) co-repressor [48]. Twist1/2 belongs to the basic-helix-loop-helix (bHLH) family of transcription factors and interact with E boxes by its bHLH domain and repress transcription of E-cadherin in a way similar to that of Snail. Twist, Snail and Slug work synergistically, controlling an overlapping and distinct sets of genes, during tumor progression. Apart from its role as a transcriptional repressor, Twist also serves as a transcriptional activator by recruiting BRD4 to coordinate EMT induction [9]. These transcription factors also regulate genes extensively involved in motility, proliferation, differentiation and survival [25, 49, 50].

Post-transcriptional Regulation

Since the discovery of ribonucleoproteins (RNP), interest in RNA biology has escalated rapidly. Given the intimate connection from transcription to translation in eukaryotic gene regulation, many studies have suggested that post-transcriptional events involving collections of mRNAs are tightly coordinated [51]. Post-transcriptional regulation of gene expression is emerging as a critical factor for many cellular and developmental processes. In many cases, RNA-binding proteins (RBPs) are the major

players by interacting with either coding or untranslated regions of mRNA [52]. RBPs are involved in every step of RNA metabolism in terms of post-transcriptional regulatory processes. The subcellular localization and level at which an (RBP) is expressed is crucial in determining its function. Nuclear RBPs primarily regulate nascent mRNA (pre-mRNA) processing events, including capping, splicing, 3'-end cleavage, polyadenylation and nuclear export. Cytoplasmic RBPs coordinate translation-associated events, such as mRNA transport, competitive or co-operative control of the translation machinery and regulation of mRNA stability [53].

1.7 Signaling Pathways that activates EMT

TGF-β Signaling

Transforming growth factor-β (TGF-β) is a prototypic member of a large family of evolutionarily-conserved cytokines that includes bone morphogenetic proteins, activins, growth differentiation factors, Nodal and inhibins. Virtually all human cell types are responsive to TGF-β, as it was evolved to regulate the expanding systems of epithelial and neural tissues, immune system and wound repair [54]. TGF-β play critical roles in embryonic development, cell differentiation, cellular homeostasis and tissue morphogenesis [54]. Mammals express three distinct TGF-β ligands (TGF-β 1-3) and three high-affinity receptors (TβR I-III). TβR-I and TβR-II both harbor Ser-Thr protein kinases in their cytoplasmic domains that are essential for intracellular signaling. TβR-III, the most abundant TGF-β receptor, lacks intrinsic enzymatic activity and modulates cellular responses to TGF-β by forming various TβR combination. In canonical TGF-β signaling, ligand binding of TGF-β to TβR-II stimulates recruitment, transphosphorylation and activation of TβR-I by TβR-II. Activated TβR-I binds and phosphorylates transcription factors Smad2/3, which forms complex with common transducer Smad4. The heteromeric

Smad2/3/4 complex accumulates in the nucleus, interact with other transcription factors and regulate gene expression in a cell- and context-specific manner [24, 55].

TGF-β also signals through several non-Smad pathways, including p38 MAPK, JNK MAPK, mTOR, RhoA, Ras, PI3K/Akt, PP2A/p70S6K, c-Src, 4E-BP1 and eEF1A1 [56]. Smad4 is essential for most but not all TGF-β-regulated transcriptional responses. TIF1γ (transcription intermediate factor 1γ, or TRIM33) is another TGF-β signal mediator that interacts with receptor-activated Smad2/3 in competition with Smad4 and engages in TGF- β-induced erythroid differentiation [57]. Most of these noncanonical TGF-β signaling pathways have been established in cell culture conditions, but their relevance to human cancer remains to be investigated.

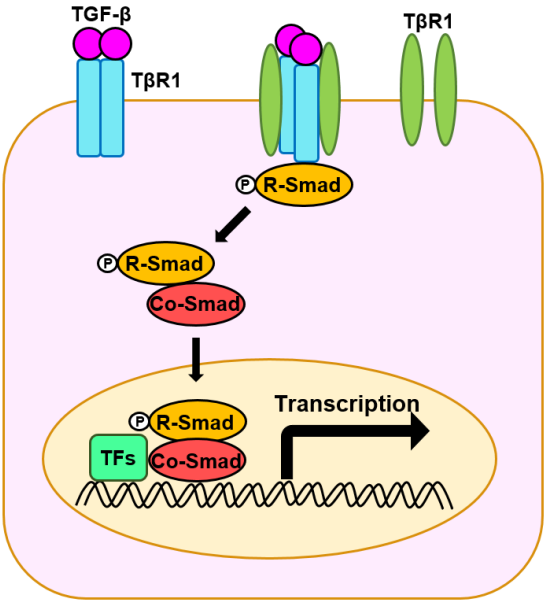


Figure 1.5 Schematic of TGF-β/SMAD-induced transcriptional responses
Activation of the pathway is initiated by binding of extracellular TGF-β ligands to the transmembrane TβR2 (TGF-β receptor 2), leading to recruitment and oligomerization of TβR1. Activated TβR1 phosphorylates intracellular mediator Smad2 or Smad3, termed regulatory-SMADs (R-Smads). R-Smads relay the signal to the cytoplasmic effector Co-Smad (Smad4, etc.), which enables translocation of the R-Smad-co-Smad complex into the nucleus and recruitment of additional co-transcriptional activators, repressors and/or co-factors to regulate expression of target genes. (TFs indicate Transcription Factors; 'P' indicates phosphorylation; Double wavy lines indicate DNA)

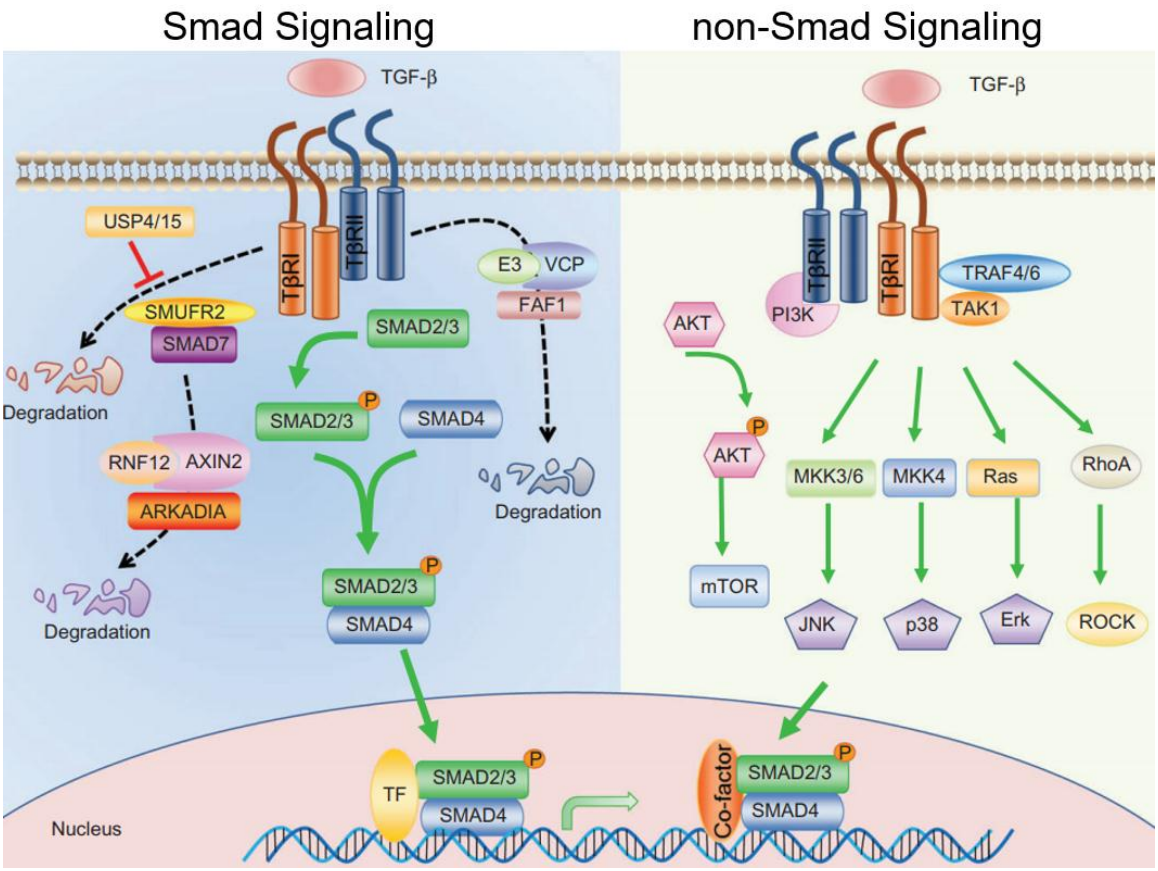


Figure 1.6 Overview of Smad and non-Smad arms of TGF-β signaling

TGF-β signals through specific TGF-β Ser/Thr kinase receptors (TβR1 or TβR2). Activated TβR1 induces Smad2/3 phosphorylation, leading to hetero-oligomerization of p-Smad2/3-Smad4 and translocation of the complex to the nucleus to regulate target gene expression. Smad7 is one of the TGF-β target genes that functions as an inhibitory Smad by recruiting E3 ligase SMURF1/2 to TβR1. ARKADIA-RNF12-AXIN2 enhances TGF-β signaling by targeting Smad7 for polyubiquitination and degradation [58, 59]. In addition, USP4/15 deubiquitinases can remove ubiquitin chains from TβR1 to stabilize the TβR1 receptor. FAF1 targets TβR2 for degradation by recruiting the VCP/E3 ligase complex, thereby limiting excessive TGF-β response. Non-Smad signaling include several branchs: PI3K-Akt-mTOR pathway, ERK, p38 and JNK MAPK cascade, and pathways downstream of Rho-like GTPase signaling intermediates [60, 61]. Both TβR1 and TβR2 are directly involved inthe activation of PI3K-Akt pathway by interacting with p85 sub-unit of PI3K [62]. Activation of Akt by PI3K induces mTOR, which controls translational responses. TGF-β-mediated activation of JNK and p38 MAPK pathway is partly mediated through ubiquitin ligase tumor necrosis factor receptor-associated factors 4 and 6 (TRAF4/6). TGF-β-stimulated interaction of TRAF4 with TβR1 induces K63-polyubiquitination of TRAF4 and, as a consequence, activation of TAK1, a MAPKKK family member. Additionally, TRAF6-TβR1 interaction induces auto-ubiquitination of TRAF6 and Lys63-linked polyubiquitination of TAK1 [63]. Concomitant activation of TAK1 results in activation of its downstream target mitogen-activated protein kinase 3,4 and 6 (MKK3/4/6), leading to

subsequent activation of p38 MAPK and JNK. In response to TGF-β, autophosphorylation of TβR1 and TβR2 triggers the recruitment of Grb2 and SOS, which activates Ras and Raf-MEK-Erk MAPK cascade. GTPase RhoA and its target ROCK are also induced to promote actin stress fiber formation and mesenchymal characteristics [64-66]. (Figure adapted from Feng Xie, et al. [66])

Under physiological settings, the anti-proliferative actions of TGF-β counteract the effects posed by local mitogenic stimulation. Under intense mitogenic stimulation, however, the TGF-β pathway triggers cytostasis or apoptosis, denpending on the intensity of the proliferatife signals, to offset increased cell proliferation. TGF-β prevents cell cycle entry into S-phase by upregulation of cyclin-dependent kinase inhibitors (such as p15 and p21) and suppresion of c-Myc [67]. Conversely, TGF-β-induced apoptosis include several Smad-dependent and –independent mechanisms depending on the cellular context. TGF-β preferentially drives differentiation of mesenchymal precursors towards fibroblasts [68]. In addition, TGF-β also suppresses cell proliferation and tumor formation by blocking the production of paracrine factors in stromal fibroblasts and inflammatory cells [54]. Collectively, TGF-β prevents tumor progression by regulating not only cell proliferation, differentiaion and survival, but also the cellular microenvironment .

TGF-β signaling is implicated as the primary inducer of EMT. In normal epithelial cells and early tumorigenic cells, TGF-β signaling inhibits uncontrolled growth by inducing cell cycle arrest and apoptosis. However, during advance stage of tumorigenesis, TGF-β ligands are frequently augmented, which may be formed by tumor cells or tumor-associated immune and stromal cells in the microenvironment of basal-like breast tumors [13, 69]. Malignant cells evade TGF-β suppressive effects either through mutational inactivation of core compoenents of the TGF-β pathway or just by disabling the tumor suppresive arm of the pathway, allowing cancer cells to freely usurp the remaining TGF-β circuit to their advantage. Particularly, breast cancer seem to preferentially disable the tumor-suppressive arm of TGF-β to benefit from tumor-derived TGF-β by using its as a shield against antitumor immunity. Collectively, TGF-β pathway facilitates cancer

progression and metastasis by inducing EMT, inhibiting immunosurveillance, activating fibroblasts and neoangiogenesis. This dramatic conversion in TGF-β function is recognized as the 'TGF-β paradox' and its detailed molecular mechanisms have not been elucidated entirely.

At the transcriptional level, TGF-β directly or indirectly activates EMT by upregulation of transcription factors, including Snail and EF1, leading to repression of epithelial marker genes (e.g. Occludins, Claudin and E-cadherin) and concomitant induction of mesenchymal markers (e.g. Vimentin, Smooth Muscle Actin and N-cadherin) [70]. TGF-β also promotes EMT by a variety of Smad-dependent and –independent effects on junction complexes Smad-induced expression of HMGA2 increases expression of Snail, Slug and Twist. TGF-β also collaborates with other signaling pathways to activate EMT and maintain the mesenchymal phenotype of metastatic tumor cells [71].

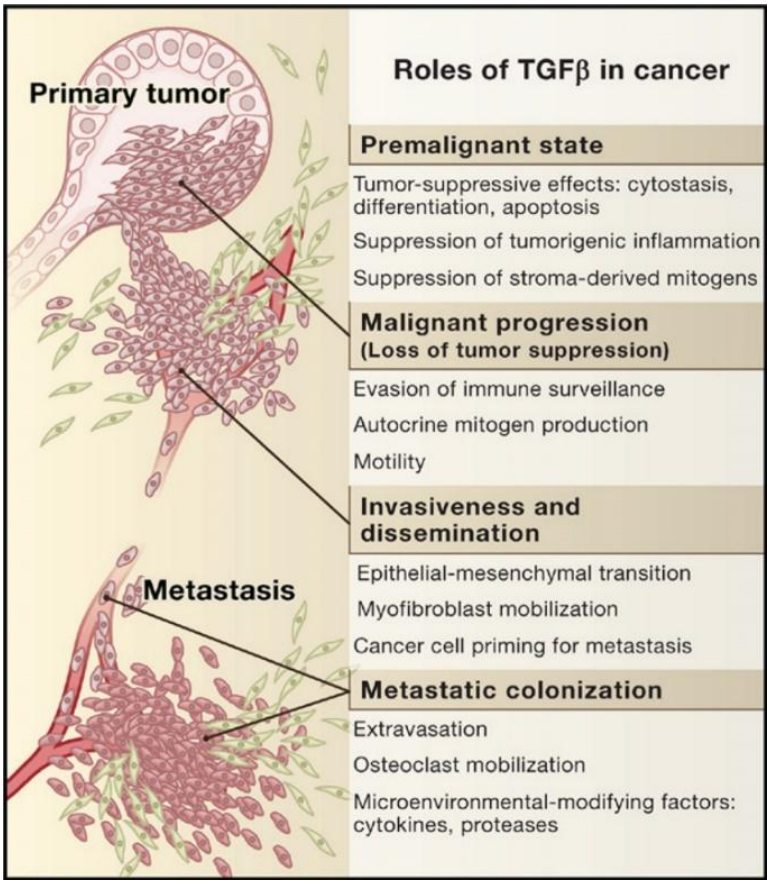


Figure 1.7 Roles of TGF-β in cancer progression

In normal epithelium and pre-malignant cells, TGF-β enforces cellular homeostasis and suppresses tumor progression directly through cell-autonomous tumor sppressive effects or indirectly through effects on the stroma. Malignant cells circumvent TGF-β suppressive effects either through mutational inactivation of core compoenents of the TGF-β pathway or by disabling the tumor suppresive arm of the pathway, allowing cancer cells to freely usurp the remaining TGF-β circuit to their advantage. Therefore, cancer cells that lose TGF-β tumor-suppressive resoponses utilize this machinery to initiate metastatic dissemination, growth factor production and immune invasion [54]. (Figure adapted from Joan M [54].)

Wnt Signaling

Three Wnt signaling pathways (canonical pathway, Wnt-calcium pathway and planar cell polarity pathway) operate in response to the binding of 19 distinct Wnt ligands to the Frizzled family of cell surface receptors [72]. Canonical Wnt pathway triggers a series of signaling events that culminate in the nuclear translocation of β-catenin, acting as a transcription co-factor, to induce the expression of a broad range of target genes involved in cell proliferation, differentiation, cell fate specification and tumorigenesis [73]. Wnt-mediated activation of the EMT program involve direct transcriptional activation of various EMT-TFs, including *TWIST*, *SNAI1*, *SNAI2*, *ZEB1*, *CDH2* and repression of *CDH1*. E-cadherin, a part of the adherens junctions complex that forms lateral connections between adjacent epithelial cells, is a negative regulator of the canonical Wnt pathway by sequestering the entire β-catenin pool at the cell membrane. In particular, inhibition of SFRP1, a Wnt antagonist, induces EMT-like changes in immortalized mammary epithelial cells TERT-HMLE and sensitizes them to TGF-β-induced EMT [74]. Additionally, TGF-β downstream target Smad7 cooperate with β-catenin to control the expression of several genes associated with cell adhesion and metastasis [75, 76].

Non-canonical Wnt signaling cascade induces EMT in a β-catenin-independent but PKC-dependent manner. Noncanonical ligands has been found to be upregulated in human epithelial cells that have completely turned into a highly mesenchymal state, compared to those that reside in an intermediate state [77].

Notch Signaling

This pathway is primarily implicated in regulating cell fate decisions, differentiation and proliferation. Four isoforms of single-pass transmembrane Notch receptors (Notch -4) are known to bind the Delta-like or Jagged family of transmembrane ligands, which triggers a series of proteolytic cleavage events that lead to production of the

active, intracellular fragment termed NICD [78, 79]. NICD translocates to the nucleus and associate with binding partners and factors, leading to expression of a cohort of target genes. NICD contains multiple domains, the RBPJ-κ association module (RAM), Ankyrin repeats (ANK domain) and transcriptional activation domain (TAD). TAD domain consists of nuclear localization sequence (NLS) and PEST domain that regulates receptor degradation [80]. PEST domain is likely to play a role in E3 ubiquitin-mediated turnover of NICD through ubiquitin-proteasome and lysosomal pathways. Mutations in the PEST domain have been demonstrated to increase the half-life of Notch 1-3 and upregulation of Notch downstream targets in TNBC [81]. During mammary gland development, Notch seems to be differentially expressed between cell subtypes. While Notch1 is higher in the luminal cells, Notch1/3 appear to mark the luminal progenitor cells. Overexpression of Numb, a negative regulator of Notch signaling by ubiquitylation and degradation of NICD, or downregulation of Cbf-1/RBPJ- κ, increases mammary stem cell (MaSC) proliferation and expands the pool of basal cells. Moreover, Notch4 is involved in promoting stem cell renewal of mammospheres [82]. Therefore, reduced Notch signaling is crucial to propagate the basal cell and MaSC population [83]. Notch pathway has also been implicated in regulating EMT in several different types of cancer [84-91]. Embryos lacking Notch1 or its partner RBPJ, fail to express Snail and thereby cannot undergo endocardial EMT [88]. Notch activates EMT by transcriptional regulation of several EMT-TFs such as *SNAI1* and *SNAI2* [92]. Numb has been shown to abrogate Notch1-mediated EMT in TNBC [93], whereas the E3 ligase MDM2 contributes to Numb degradation, leading to activation of Notch in BC. Targeting MDM2 led to a reduction in Notch signaling in MCF7 cells [94]. Notch also crosstalk with TGF-β pathway to induce EMT, via interaction of Smads with NICD and other TFs to regulate mesenchymal fate-related genes. TGF-β also can induce expression of Notch ligands such as Jagged 1, which act in an autocrine fashion [88].

Growth Factor Signaling

Binding of growth factor receptors by their cognate ligands stimulates receptor dimerization and activation of receptor-associated tyrosine kinases (RTKs). Phosphorylation of these receptors enables activation of the PI3K-Akt, ERK-MAPK, p38 MAPK and JNK pathways, promoting cell proliferation, migration and motility via induction of EMT [95]. Upon EGF stimulation, Stat3 binds to the promoter of *TWIST* in MCF7 cells. EGF was also shown to induce nuclear co-localization of Snail and phospho-Smad2/3 in MDA-MB-231 cells [96, 97]. Fibroblast growth factor (FGF) has been implicated with the induction of EMT via MAPK and MEK-ERK pathway [98, 99]. Hepatocyte growth factor (HGF), the ligand of MET tyrosine kinase receptor, is a potent activator of EMT by upregulating Sail1 [100]. Platelet-derived growth factor (PDGF) is found to provide an autocrine feedback loop to maintain the neoplastic human mammary epithelial cells in a more mesenchymal phenotype [101].

1.8 RNA Binding Protein

As soon as a gene is transcribed, many post-transcriptional events are expected to occur before the protein product is synthesized, such as mRNA processing, nucleocytoplasmic export, mRNA localization, mRNA stabilization and translational regulation. A majority of these events are controlled through a complex network of RNA/protein interactions involving recognition of specific target mRNAs by RNA binding proteins (RBPs) [102]. Post-transcriptional regulation of gene expression results from spatiotemporal equilibrium and dynamics between the regulatory sequences found on the mRNA, RBPs as well as the signaling pathways that modify them under a particular cellular context. Alteration of any of these determinants may disturb the equilibrium and thus the expression level of a given protein, causing a broader effect to cell homeostasis. Deregulation of gene expression may enable the cell to re-enter cell cycle, conferring them

growth or motility advantage over normal cells [102]. Currently, approximately 500 RNA binding domains (RBDs) have been characterized in the human genome.

While a growing body of reports published on the association of specific factors with a given RNA, these studies merely unveiled isolated pieces of jigsaw puzzle out of the entire landscape of post-transcriptional regulation. The dynamics and kinetics of the pivotal RBPs and their targets must be fully addressed in order to comprehend how to achieve a certain expression level of a given gene and how it affects oncogenesis.

1.9 Alternative Splicing

Overview

Splicing is largely carried out by the main spliceosome, a complex machinery composed of five small nuclear ribonucleoprotein particles (U1, U2, U4, U5 and U6 snRNP) and a large variety of auxiliary proteins [103]. Alternatively, a small number of introns are processed by the minor spliceosome comprised of U11, U12, U4atac/U6atac and U5RNPs [104]. The main spliceosome machinery recognizes short consensus sequences at the exon-intron junctions and catalyzes two transesterification reactions necessary for the inclusion of exons and removal of introns. Due to the short and degenerate nature of the splice sites, additional factors are often required to assist the spliceosome function. The activity of the spliceosome is strictly regulated by both *cis*-acting sequences on the pre-mRNA and transacting factors, which may either enhance or inhibit recognition of splice site and splicing reaction [105]. Two main classes of RBPs that regulate splicing by binding to the *cis*-acting elements are Ser/Arg rich (SR) proteins and the heterogeneous nuclear ribonucleopreoteins (hnRNPs), one exert positive regulation and another function as antagonistic inhibitor of splicing [106].

Alternative splicing adds an additional layer of complexity to the splicing process by the presence of exons characterized by weak element defining exon-intron boundaries.

Differential assortment of weak or variable exons allows a gene to produce multiple splice variants, encoding protein isoforms with different or even opposite function of distinct patterns of spatiotemporal expression [107]. The advent of high-throughput sequencing technologies has revealed that over 95% of human genes are estimated to undergo alternative splicing to expand proteome diversity by producing multiple mRNA and protein isoforms per gene [108]. This RNA processing event is mediated by the intricate interplay between splicing factors and defined RNA sequences/splice sites within the pre-mRNA. Pre-mRNA splicing occurs in multivariate modes, the most common of which is alternative inclusion of cassette exons. Other common modes include alternative 5' or 3' splice sites, intron retention, mutual exclusion of cassette exons, exon scrambling and trans-splicing. While tissue specificity for the dominant isoform per gene is found in more than half of all genes, switch-like events occur in about 35% of genes between any two tissue types [109].

Alternative splicing plays a pivotal role in controlling core cellular processes, such as proliferation, metabolism, apoptosis, physiological decisions, induction of differentiation and maintenance of pluripotency. Aberrant regulation of alternative splicing contributes to the onset or progression of several human diseases, including cancer [107]. As documented by numerous studies, specific splice variant signatures are strongly associated with particular types of cancer, thus representing suitable targets for the development of valuable antitumor therapies. In light of this, great interest has arisen for the investigation of the tumorigenic vs. normal splicing patterns and the search of approaches to switch splicing patterns from tumoral variant toward non-tumoral isoform. Specifically, two possibilities have been proposed: one is to target the specific activity of the oncogenic splice variant, another is to target the mechanism driving the aberrant splicing event [107].

Epithelial Splicing Regulatory Proteins and EMT

Alternative splicing must be tightly-regulated, in a spatiotemporal manner, to ensure the expression of functionally different splice isoforms. Changes in cellular phenotype, such as EMT, is modulated at the level of alternative splicing by a number of splicing factors. Alternative splicing was the first post-transcriptional mechanism linked to EMT [110]. Epithelial splicing regulatory protein (ESRP) 1 and 2 were first identified as specific regulators of fibroblast growth factor receptor-2 (FGFR2) splicing as well as other epithelial-specific variants. Cell type-specific expression of epithelial and mesenchymal isoforms of FGFR2 splicing occurs by inducing the switch of mutually exclusive exons IIIb and IIIc, respectively. Microarray analysis of PNT2 epithelial cells after ESRPs knockdown showed that loss of splicing program alone can induce some of the phenotypic changes that occur during EMT, suggesting that the ESRP-regulated splicing program is an essential aspect of the epithelial phenotype and many EMT-associated cellular changes are due to functional alterations of proteins that undergo isoform switch during this process.

Role of Alternative Splicing in Breast Cancer

Aberrant alternative splicing (AS) events are linked to BC onset and progression. Carriers of *BRCA1* mutations are predisposed to breast cancer with lifetime risk of up to 80% than non-carriers [111]. Mutation of the *BRCA1* gene within specific sequences affects the binding of splicing factors, leading to AS events of *BRCA1* and production of variants that lack functional domains of the protein, thus compromising its tumor suppressor activity [112, 113]. Similarly, point mutations may also occur in splicing factors, resulting in inaccurate AS of cancer-related genes, such as favoring proto-oncogene splice variants over tumor suppressor splice variants [114]. Deregulation of splicing factors, and also other RBPs, can also cause repercussions on splice site selection, and thereby are associated with development and progression of BC (Figure) [112].

1.10 mRNA Stability and Translation

The majority of mRNA regulatory elements involved in modulation of post-transcriptional events are situated within the 5' and 3'-UTR, where they act as platforms for the assembly of regulatory factors. While the 5'-UTR is primarily engaged in controlling mRNA translation, the 3'-UTR regulate multiple aspects of the mRNA metabolism, such as nuclear export, cytoplasmic localization, mRNA stability and translational efficiency [115-117].

Tight regulation of mRNA half-life plays a pivotal role in normal cell functions. Substantial stability renders a mRNA available for translation for a longer time, yielding high levels of protein products. Differences in the length and structure of the 3'-UTR expands the mammalian gene products generated by AS and alternative polyadenylation [118, 119].

The structure of an mRNA- 5' untranslated region (5'-UTR), 5' cap structure, open reading frame (ORF), 3' untranslated region (3'-UTR) and 3' terminal poly (A) tail-governs mRNA stability and translational efficiency. RNA sequence elements like 5' cap and 3' poly(A) tail are universally present in all mRNAs and convey constitutive processes without apparent selectivity of one mRNA relative to another [109]. However, numerous *cis* elements have been described which pose effects on stability and/translation of given subsets of mRNAs. In the 5'-UTR, the iron-response elements, JNK-response elements and turnover determinants are present in the chemokine ligand 1 (KC) mRNA. These elements dictate the activities of *trans* factors which elicit responses on several processes such as translation, turnover, storage and transport. mRNA decay elements such as CRD-1 have been observed in the ORF of mRNAs such as *c-fos*, *c-myc* and *β-tubulin*. 3'-UTRs are well recognized to contain turnover and translation determinants that bind RBPs to target specific mRNA for stabilization/destabilization or translational activation/repression. To-date the most commonly found 3'-UTR *cis* element are the AU-

rich elements (AREs), which often contain a variable number of AUUA pentamers, sometimes harbored within a U-rich region. AREs are found on numerous mRNAs encoding oncogenes, cytokines, interleukins, TNF-α and cell-cycle regulators such as *c-fos*, *c-myc* and *cyclins A*, *B1* and *D1*. Many of them are overexpressed during cellular transformation due to mRNA stabilization or enhance translation [102, 109]. Other specific secondary structures such as stem-loop motifs are found on 3'-UTR of cell cycle-regulated histone mRNAs [109]. Together, these regulatory elements serve as binding sites for a variety of RBPs that modulate mRNA stability and translation efficiency. Given the abovementioned involvement of these regulatory elements and trans-acting factors (e.g. RBPs), alterations in any of these components can cause major impact on mRNA half-life and/or translation, resulting in aberrant levels of expressed protein and hence metabolic changes leading to disease.

Regulation of mRNA stability and translation occurs via interaction of *cis* elements with *trans*-acting factors, which in turn target the mRNA for rapid degradation or protect it from nuclease access and/or regulate translational efficiency.

Several *trans*-acting factors are emerging as core regulators of expression of cancer-related genes. In general, the levels of RBPs are frequently elevated in cancer. Each RBP is likely to regulate a discrete but broad subset of target transcripts simultaneously, thus leading to an expanding functional network of changes which pose significant consequences for cancer cell biology [120]. Some of the changes occur at the level of AS, generating variants that promote multiple aspects of tumorigenesis, which enables cancer cells to rapidly adapt to adverse conditions encountered during transformation, leading to chemoresistance.

Cancer genes are characterized by their altered gene expression and/or activity leading to abnormal phenotype. Such changes confer the cell with competitive growth advantages: enhanced cell division, resistance to cell death, increased angiogenesis, invasion and metastasis, and evasion of immunosurveillance. Apart from gene mutation,

additional mechanisms also contribute to abnormal level of gene products in cancers, such as gene dosage, gene transcription, post-transcriptional control of the mRNA and regulated proteolysis [109].

RBPs that Regulate mRNA Stability and/or Translation

RBPs influence translational efficiency primarily through three modes of action: direct interaction with mRNA, bridging other RBP from its target mRNA, and tag a mRNA for rapid deadenylation/degradation or to protect it from nucleases [121].

One of the best studied mRNA stabilizing protein is HuR, a ubiquitously expressed member of the embryonic lethal abnormal vision (ELAV) family of ribonucleoproteins. HuR regulates cyclin A and B1 mRNA stability in a cell cycle-dependent fashion by nuclear-cytoplasmic shuttling of HuR. Other members of this family, such as neuronal specific HuB, HuC and HuD, also participate in mRNA stabilization. HuR is well known of its ability to recognize AU-rich elements (AREs) found on 3'-UTR of specific mRNAs, which influences different cellular processes such as proliferation, differentiation, apoptosis, inflammation, stress response and cancer [51, 122]. Snail mRNA is one of the HuR targets. The binding of HuR to Snail 3'-UTR is significantly increased following exposure to H_2O_2, which triggers EMT in several cell types [123]. HuR has been reported to bind to p53 3'-UTR to enhance its translation [124].

CUG triplet repeat RNA binding protein 1 (CUGBP1), or ELAV-like family member 1, is another example of mRNA stabilizing protein. CUGBP1 and CUGBP2 are members of the CELF (CUGBP and ETR3-like factors) family of RNA binding proteins. Upon binding to the ARE in the cyclooxygenase-II (COX-2) mRNA, CUGBP stabilizes COX-2 mRNA to repress its translation [102].

Several RBPs have opposite effects on mRNA stability. AU-rich element RNA binding protein 1 (AUF1), or hnRNP D, consists of four isoforms of 37, 40, 42 and 45 kDa.

AUF1 is correlated with rapid degradation of ARE-containing mRNAs [125]. Tristetraprolin (TTP or ZFP36), characterized by two tandem repeat zinc finger motifs, binds to AREs and mediate mRNA decay [126]. Several other targets of TTP include TNF-α, VEGF, IL-1, IL-8, GM-CSF and HIF-1 [127]. KH-type splicing regulatory protein (KSRP) is another example of mRNA decay protein by binding to AREs [128].

1.11 Introduction to RBMS3

RNA binding motif, single-stranded interacting protein 3 (RBMS3) belongs to the small family of *c-myc* single stranded binding proteins, which contains two ribonucleoprotein domains (RRM). RBMS3 was initially identified as binding to an upstream element of the mouse collagen α2 gene promoter [129]. RBMS3 was later shown as a tumor suppressor by regulating G1/S progression, cell proliferation and inhibit angiogenesis in gastric cancer and nasopharyngeal carcinoma [130-132]. Downregulation of RBMS3 was also found to facilitate development and progression of lung squamous cell carcinoma. RBMS3 was to be expressed in activated hepatic stellate cells, a type of mesenchymal cells of the liver, and liver fibrosis [133]. Despite this, the effect of RBMS3 appears to be controversial. Recent literature suggested that loss of RBMS3 confers chemoresistance to epithelial ovarian cancer via activation of miR-126/β-catenin/CBP signaling [134]. RBMS3 has been reported to function by binding to the 3'-UTR of its targets to increase mRNA stability and half-lives in gastric cells, activated hepatic stellate cells and zebrafish embryo [133, 135, 136].

One of the best developmental function of RBMS3 is studied in zebrafish. Zebrafish *RBMS3* was transiently expressed in the cytoplasm of condensing neural crest cells within the pharyngeal arches. Morphants for *RBMS3* demonstrated severe craniofacial defect phenotype resembling cartilage/crest defects observed in *Tgf-βr2:Wnt1-Cre* mutants, with reduced proliferation of prechondrogenic crest and significantly altered expresison for

chondrogenic/osteogenic lineage markers. Zebrafish RBMS3 posttranscriptionally regulate one of the major cartilage differentiation effectors, the TGF-βr pathway, to driving central neural cells down to a chondrogenic lineage [136].

CHAPTER 2. RATIONALE, SPECIFIC AIMS AND INNOVATION

Rationale and Specific Aims

Triple-negative breast cancer (TNCB) represents up to 15-20% of all breast cancers and show enhanced invasiveness, metastatic potential and worse prognosis. Metastasis, the cause of 95% of cancer-related deaths, represents the biggest clinical challenge accounting for the vast majority of cancer-related deaths. However, effective therapeutic strategies targeting breast cancer metastasis are still scarce due to spatiotemporal intra-tumor cellular heterogeneity. Improved approaches to identify TNBC in the clinic and better understanding of the molecular programs that defines the metastatic potential of TNBC are urgent.

Cancer metastasis is mediated by cellular interactions in response to signals from the tumor microenvironment affecting dynamic remodeling of the actin cytoskeleton contributing to the modulation of cell adhesion, migration and invasion. TGF-β is known to be highly expressed in the breast tumor microenvironment. TGF-β acts as a negative growth factor via growth inhibition and apoptotic induction in healthy cells and early tumor cells, however, it switches to promoting cell invasion, angiogenesis and cell adhesion and migration at later stages of tumorigenesis [137]. While the mechanism underlying the switch from a growth suppressor to a metastasis promoter is largely unknown, understanding how cancer cells interpret TGF-β signals from the microenvironment is necessary for elucidating the process. Previous evidence from zebrafish demonstrated the physical interaction between RBMS3 and Smad2/3 transcripts, establishing the role of RBMS3 in regulating chondrogenesis by stabilizing the pool of Smad2/3 transcripts in order to maintain a high level of TGF-β signaling and thus cells remain at the mesenchymal state. Since RBMS3 is highly expressed in BLBC cells and its function in BC is unknown, here we propose BLBC cells exploit the same

mechanism to fuel the TGF-β pathway. Hence, the **specific aims** of the first part of my work are: 1) to study the function of RBMS3 in BC cell migration and invasion; 2) to investigate the role of RBMS3 in TGF-β signaling.

90% of BC metastasis occurs in the mammary duct. Although metastatic BC cells often express basal markers, the basal/myoepithelial layer is not thought to be the source of metastasis, but rather a barrier for metastasis. Therefore, it becomes an enigma how the organized luminal cells break this entity and spread to other organs. At the posttranscriptional level, the identity of luminal cells is strictly governed by lineage determinants like the master regulator ESRP1/2 proteins. ESRPs not only organize cell polarity, cell migration and proliferation, but are also known to suppress the expression of EMT genes. Unlike the less metastatic luminal BC cells that overexpress ESRPs, metastatic TNBC cells express little to no ESRPs. Conversely, in our preliminary work, we found that TNBC cells preferentially express an RBP, RBMS3, which is absent in luminal BC cells. The mutually exclusive pattern of ESRPs and RBMS3 led us to speculate that RBMS3 is a mesenchymal lineage determinant that may be required for enhancing or maintaining mesenchymal phenotypes. Therefore, we hypothesize that RBMS3 is a tumor promoter in the TNBC cells. We also speculate that if we swap the ESRPs with RBMS3, luminal BC cells may lose their epithelial identity and adopt mesenchymal features. Therefore, the **specific aims** of the second part of my work are: 1) to understand the functional similarities and differences of ESRP1, ESRP2 and RBMS3; 2) to perform RNA-seq analysis on the established cells and study the posttranscriptional profiles to understand the networks controlled by these RBPs; and 3) to identify common targets that may be critical for BC cell migration, adhesion and invasion.

Overall, my **hypothesis** is that RBMS3 is a tumor promoter involved in regulating RNA metabolism and splicing in TNBC.

Innovation

The study is based on the hypothesis that RBMS3 plays tumor-promoting roles through its abilities to regulate RNA metabolism or splicing. The overall innovations are: 1) This is the first study to demonstrate the oncogenic function of RBMS3, which will provide a better understanding to the impact of RBMS3 on BLBCs. 2) This is also the first to delineate RBMS3-mediated regulation of RNA metabolism and splicing in BC. 3) This study proposes a novel strategy to study three RBPs by swapping ESRPs with RBMS3 in luminal BC cells.

CHAPTER 3. MATERIALS AND METHODS

3.1 Cell Lines and Cell Culture

The MCF7, MDA-MB-231, MDA-MB-157 and Hs578T breast cancer cell lines were grown in Dulbecco's modified Eagle's medium (DMEM)/F12 supplemental with 10% fetal bovine serum (FBS), 100μg/mL streptomycin and 100 unit/mL penicillin. T47D and BT-549 breast cancer cell lines were grown in RPMI1640 plus 10% FBS. All cells were cultured at 37℃ in humidified air containing 5% CO_2. For establishing stable clones of RBMS3 knockout or overexpression, transfected cell lines were selected with puromycin (1 μg/mL) for 4 weeks followed by selection of single clones.

3.2 Plasmids, Drugs, Antibodies and Primers

Stable RBMS3 knockout cells were generated by lentiCRISPR V2 purchased from Addgene. LentiCRISPR V2 vector was digested with *BsmB*I and ligated with indicated annealed oligonucleotides. Targeting sequence for RBMS3 is as the following: CAGCTACATGGGCAAACGCC. Stable ESRP1 and ESRP2 knockout cells were generated by lentiCas9-Blast purchased from Addgene. Targeting sequences for ESRP1 and ESRP2 are CTGGACCAAGCCCTCCGA and AGTCGGTCTCGTCCGAGCC, respectively.

Anti-ESRP1 antibody (HPA023719) was from Sigma. Anti-ESRP2 antibody (ab155227 and ab113486) was from Abcam. Anti-RBMS3 antibody (GTX47423) was from Genetex. Antibodies against E-cadherin (14472), Snail (3879S), Twist (46702S), Slug (9585S) and Vimentin (5741T) were from Cell Signaling. Anti-Smad2/3 antibody was a sample from Santa Cruz. Anti-Smad2, Smad3 and Smad4 antibodies were sample kit from Cell Signaling. Anti-β-actin antibody (MA5-15739) was from Thermo. Anti-Strep-Tag antibody (2-1509-001) and Strep-Tactin XT Superflow 50% suspension (2-4010-002) were from IBA Life Science. Recombinant human TGF-β1 was purchased from Peprotech (100-21). A8301 (SML0788) and actinomycin D (A1410) were purchased from Sigma-Aldrich.

Table 1 Primers for qPCR

Gene	Forward Primer	Reverse Primer
SMAD2	AGCAGGAATTGAGCCACAGAGT	AAGAGTAGTAGGAGATAGTTCT
SMAD3	ACCACTACCAGAGAGTAGAGA	TGGGGCTCGATGCCTGCGGGGA
SMAD4	ACTGCCAACTTTCCCAACATT	ACCAGTAAATCCATTCTGCTGCT
SMAD7	CCCCATCACCTTAGCCGACTCTGC	CCCAGGGGCCAGATAATTCGTTCC
CTGF	ACTGTCCCGGAGACAATGAC	TGCTCCTAAAGCCACACCTT
PAI1	ATTCAAGCAGCTATGGGATTCAA	CTGGACGAAGATCGCGTCTG
ID2	TCAGCCTGCATCACCAGAGA	CTGCAAGGACAGGATGCTGAT
MMP1	ATCGGGGCTTTGATGTACCC	GGCTGGACAGGATTTTGGGA
PTGS2	AGTCCCTGAGCATCTACGGT	GCCTGCTTGTCTGGAACAAC
OCT4	GGAGGAAGCTGACAACAATGAAA	GGCCTGCACGAGGGTTT
SOX2	TGCGAGCGCTGCACAT	TCATGAGCGTCTTGGTTTTCC
BMI1	TGCTGGAGAACTGGAAAGTG	GATGAGGAGACTGCACTGGA
β-Actin	AGAGCTAGCTGCCTGAC	GGATGCCACAGGACTCCA

3.3 Cell Viability Assay

This assay was performed essentially described by CellTiter-Glo 2.0 Manual. Briefly, luminal cell lines (T47D and MCF7) were plated at a density of 20,000 cells per 96 well plate (in triplicate) and allowed to grow for a specific time course. At desired time point, add reagent equal to the volume of medium in each well. Allow mixing for 2 minutes on orbital shaker and another 10 minutes of incubation at room temperature to induce cell lysis and stabilization of signals. Luminescent signals were measured on Synergy HTX Multi-Mode Microplate Reader.

3.4 Transwell Migration and Invasion Assay

Cell were grown to 80% confluency, resuspended in serum-free medium and seeded into the upper chamber of Transwell inserts with an 8 μm pore size membrane (Falcon) at $5x10^4$ cells per insert. For the invasion assay, upper chambers of Transwell inserts were pre-coated with 50 μL of Matrigel (BD Biosciences), which was allowed to dry for 1 hour prior to seeding of cells for invasion. The cells were allowed to migrate toward DMEM/F12 supplemented with chemoattractant epidermal growth factor (EGF) (20ng/mL) in the lower chamber. Basal-like MDA-MB-231 and BT-549 cell lines were

incubated for 8 hours for migration assay and 24 hours for invasion assay. T47D and MCF7 cells were incubated for 24 hours for migration assay and 48 hours for invasion assay. At the end of incubation period, cells remaining in the upper side of the Transwell insert membrane were mechanically removed using a cotton swab. The cells on the bottom of the Transwell insert were washed twice with PBS, fixed with 4% paraformaldehyde (2 minutes), methanol (10 minutes), washed once with PBS, and then stained with 0.5% (w/v) crystal violet in 20% methanol (v/v) (10 minutes). The stained cells in 4 non-overlapping fields from each insert were counted with an inverted microscope, using the 20X objective (Eclipse TS100, Nikon Instruments, Melville, NY). All experiments were performed in triplicates.

3.5 Colony formation Assay

Colony formation assay was performed using double-layer soft agar in 6-well plates with a top layer of 0.35% agar and a bottom layer of 0.7% agar. Briefly, cells were suspended in 0.35% agar medium and laid on the top of the supporting agar layer (0.7% agar) in 6-well plates, and fed twice per week. Colonies were allowed to form in an incubator at 37 ℃ and 5% CO_2 for 14 days. At the end of the incubation time, cell cultures were photographed and the colonies were stained and counted.

3.6 Would Healing Assay

Cells for scratch assay were seeded at same density on 6-cm dish. At about 90% confluency, cells were starved overnight. Next morning, a consistent wound (gap) was created using pipette tip. After removal of cell debris with PBS, cells were replaced with fresh medium and incubated for 24 hours. Gap measurement was performed by ImageJ software. Migration rates were calculated using data from 10 measurements of random gap points.

3.7 Mammosphere Assay

Mammosphere assay was performed following the protocol previously described [138]. Briefly, cells were seeded in single-cell suspension in triplicates into ultra-low

attachment 6-well plates (Corning) in DMEM/F12 medium supplemented with 20 ng/mL EGF, 5 mg/mL insulin, 0.5 mg/mL hydrocortisone and 2% B27. After 2 weeks of incubation, the presence of spheres was assessed by inverted microscopy. At least 20 random fields for each cell lines were visualized; the number and size of spheres in the 20 fields were calculated as a percentage over that of parent cells.

3.8 3D On-top Matrigel Assay

This assay was performed following an established protocol. Essentially, pre-chilled culture surface was coated with a thin layer of Matrigel. Single cell suspension was subjected to centrifugation at ~115 g, resuspended in half the medium volume and plated onto coated surface at $0.2x10^5$ /cm^2. A layer of 10% Matrigel (in medium) was added on top to stabilize the cells and culture was maintained for 4 days.

3.9 Anchorage-independent Growth Assay

To observe the adhesion-independent growth of 14 luminal BC-derived clones, $0.8x10^3$ cells were seeded in 96-well round bottom ultralow-attachment plates. Cells were cultured for up to one week and photographs were taken daily.

3.10 RNA Isolation and Quantitative Real-Time PCR (qPCR)

Total RNA was prepared using the RNeasy Mini kit (Qiagen) according to the manufacturer's instructions. Specific quantitative real-time PCR experiments were performed using SYBR Green Power Master Mix following manufacturer's protocol (Applied Biosystems). All values were normalized to the level of β-actin.

3.11 Immunofluorescence Staining and Western Blot Analysis

Experiments were performed as described previously [139, 140]. For immunofluorescence staining, cells were grown on cover slips, fixed with 4% paraformaldehyde and incubated overnight with anti-Strep-Tag primary antibodies. Secondary antibodies were 568 goat anti-mouse IgG (H + L) (Molecular Probe, Carlsbad, CA). Finally, cover slips were visualized under Nikon confocal microscope.

For Western blot analysis, cells were rinsed and collected on ice with cold PBS by scraping. After centrifugation at 1000rpm for 3 minutes and removal of PBS, cell pellet was lysed with lysis buffer (50mM Tris-HCl pH 7.4, 150mM NaCl, 0.2mM EDTA, 1% Triton-X100, 10% glycerol, 5mM β-glycerophosphate, 5mM 4NPP, 1mM Na_3VO_4, 5mM NaF, 2 μg/mL aprotinin, 1 μg/mL leupeptin and 0.5 tablet of protease inhibitor cocktail). Cells were homogenized by gentle up-and-down pipetting and lysed for 10 minutes on ice followed by sonication for 1 minute. After high speed centrifugation for 15 minutes, supernatants were transferred in new tubes and protein levels were quantified via Bradford assay followed by normalization. Sample buffer (4X) was added at a ratio of 1:3 and proteins were then separated by SDS-PAGE gels, transferred to PVDF or NC membrane and immunoblotted.

3.12 Crosslinking RNA Immunoprecipitation (CLIP)

This experiment was performed based on an established protocol. Cells were grown on 15 cm dish and cultured to around 90% confluency. To crosslink in vivo, 37% formaldehyde was added directly to cells in medium dropwise to a final concentration of 0.75% and incubated with gentle shaking for 10 minutes at room temperature. 125mM glycine was then added for 5 minutes. Cells were rinsed and collected on ice with cold PBS by scraping. After centrifuge at 1000rpm for 3 minutes, cell pellet was resuspended with equal volume of polysome lysis buffer (100 mM KCl, 5 mM $MgCl_2$, 10 mM HEPES pH7.0, 0.5% NP40, 1 mM DTT, 100 unit/mL RNase Out, 400 μM VRC, 10 μL protease inhibitor cocktail) by gentle up-and-down pipetting. After high speed centrifuge for 15 minutes, cell lysate was transferred to new tube and quantified for protein level. 20 μg lysate was saved as RNA input at -20 ℃ for later use. Meanwhile, 100 μL Strep-Tactin beads were washed twice with ice-cold NT2 buffer (50 mM Tris-HCl pH7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.05% NP40). After final wash, beads were resuspended in 850 μL ice-cold NT2 buffer supplemented with 200 units of RNase Out, 400 μM VRC, 1 μL of 1M DTT and EDTA to 20 mM. For each RIP reaction, 2 mg cell lysate was added to the beads and the total reaction volume was adjusted to 1 mL. Bead-lysate mixture was incubated on rotation for 3 hours at 4 ℃. Beads were pelleted by centrifuge at 2000rpm for 3 minutes at 4 ℃ and washed four times with ice-cold NT2 buffer with rotation at 4 ℃. Meanwhile,

previously-collected cell lysate (RNA input tube) was thawed on ice. At final wash, 1/5 beads were collected for western blot analysis. After final wash, 1 mL Trizol reagent (Invitrogen) was added directly to the beads and cell lysate (RNA input tube). RNA was extracted following the manufacturer's manual. 30 μg glycogen was added as a carrier to aid in RNA precipitation. After cDNA synthesis, quantitative real-time PCR was performed to determine the levels of desired transcripts. All equipment was kept at RNase-free level and all reagents were prepared with RNase-DNase-free H_2O.

3.13 Fluorescence-Activated Cell Sorting (FACS)

Cell Surface Marker Assay: Cells were detached from plates, blocked with 2% BSA for 30 minutes and incubated with anti-human CD24 (PE-conjugated, ebioscience), anti-human CD44 (PE-Cy7-conjugated, ebioscience), anti-human CD49f (PE-Cy7-conjugated, ebioscience) or EPCAM and finally analyzed using FACSCalibur flow cytometer.

Annexin Early-Apoptosis Assay: Experiment was performed essentially as described by manufacture's protocol using Annexin V-FITC Early Apoptosis Detection Kit (Cell Signaling Technology). Briefly, MCF7 cells stably expressed control or RBMS3 vector were treated with TGF-β1. 10^5-10^6 cells were detached from plates and resuspended with ice-cold Annexin V binding buffer. 1 μL Annexin V-FITC conjugate and 12.5 μL Propidium Iodide (PI) solution was added to 96 μL cell suspension. After 10 minutes of incubation on ice in the dark, cell suspension was diluted to a final volume of 250 μl/assay with ice-cold Annexin V binding buffer. Analysis was performed immediately using FACSCalibur flow cytometer.

Cell cycle analysis: The experiment was performed essentially as previously described [141]. Briefly, cells were seeded and treated with TGF-β1 for 24 hours. Cells were harvested and fixed in cold 70% ethanol at -20 ℃ for overnight. Cells were washed with phosphate buffered saline (PBS) twice and stained by 50 μg/mL propidium iodide

(PI) solution with 100 µg/mL RNase A in dark for 30 minutes. Cells were analyzed using a FACSCalibur flow cytometer.

3.14 Drug Treatment

mRNA stability assay was accomplished by inhibition of global RNA synthesis through the use of pharmacological inhibitor Actinomycin D (ActD). Cells were treated with ActD at 5 ng/mL for 0, 3, 6 and 9 hours. At the end of incubation time point, medium was removed and cells were directly lysed on dish with Trizol reagent.

3.15 Trypan Blue Exclusion Assay

This assay was performed to determine cell death. Briefly, cells were left untreated or treated with TGF-β1 at 10 ng/mL for indicated periods of time. At the end of incubation time point, all cells (both adherent and floating) were harvested, stained with trypan blue and counted immediately to determine the percentage of nonviable cells (stained blue). Columns represent the mean of three independent measurements and bars represent standard error.

3.16 RNA-sequencing (RNA-seq)

Total RNAs isolated from T47D, T47D-RBMS3, T47D-ESRP1 KO, T47D-ESRP2 KO, T47D-ESRP DKO and T47D-ESRP DKO-RBMS3 cells using RNeasy Mini Kit (Qiagen). RNA integrity and concentration were determined using Bioanalyzer (BioTek) and agarose gel electrophoresis. The purified samples were sequenced and the clean reads were mapped against human genome. The normalized expression values for each gene were quantified using $\log_2$FPKM value (expected number of fragments per kilo base of transcript sequence per million base pairs sequenced). The transcript levels of genes having a fold change (FC) of greater than 0.5 were significantly differential between two samples. The software Morpheus(https://software.broadinstitute.org/morpheus/) was used to draw heatmap.

3.17 Functional Enrichment Analysis for DEGs

DEG sets were screened out for functional enrichment analysis. Essentially, Gene Ontology (GO) terms for biological processes, cellular components and molecular function categories were annotated by two public resources: GO (http://geneontology.org/) and Metascape (http://metascape.org/) [142-144]. Only terms with p-value <0.01 and the number of enriched genes greater than 3 were considered as significant. All resultant terms were then grouped into clusters based on their similarities. The most enriched term within a cluster was chosen to represent the cluster. DEG sets were also used to conduct gene set enrichment analysis (GSEA), which was performed with GSEA v3.0 on various functional characteristic gene signatures. Gene sets were obtained from Molecular Signature Database (MSigDb).

3.18 Breast Cancer Survival Analysis

Kaplan-Meier survival analyses for clinical outcomes (RFS or DMFS) of breast cancers were performed using web tool Kaplan-Meier Plotter (http://kmplot.com/analysis/) and UCSC Xena tool (https://xenabrowser.net/). The percentiles of the patients between the upper and lower quartile were auto-selected based on the best performing thresholds as cutoffs.

3.19 Correlation analysis

Correlation analysis was analyzed with the Gene Expression Profiling Interactive Analysis (GEPIA) online analysis tool, which uses the BRCA dataset from TCGA database (http://gepia2.cancer-pku.cn/).

3.20 Human Breast Cancer Tissue Dataset Analysis

mRNA expression analysis of breast tumor datasets was conducted on web tool Oncomine (https://www.oncomine.org/). Correlation analysis was analyzed with the Gene Expression Profiling Interactive Analysis (GEPIA) online analysis tool, which uses the BRCA dataset from TCGA database. (http://gepia2.cancer-pku.cn/).

3.21 Statistical analysis

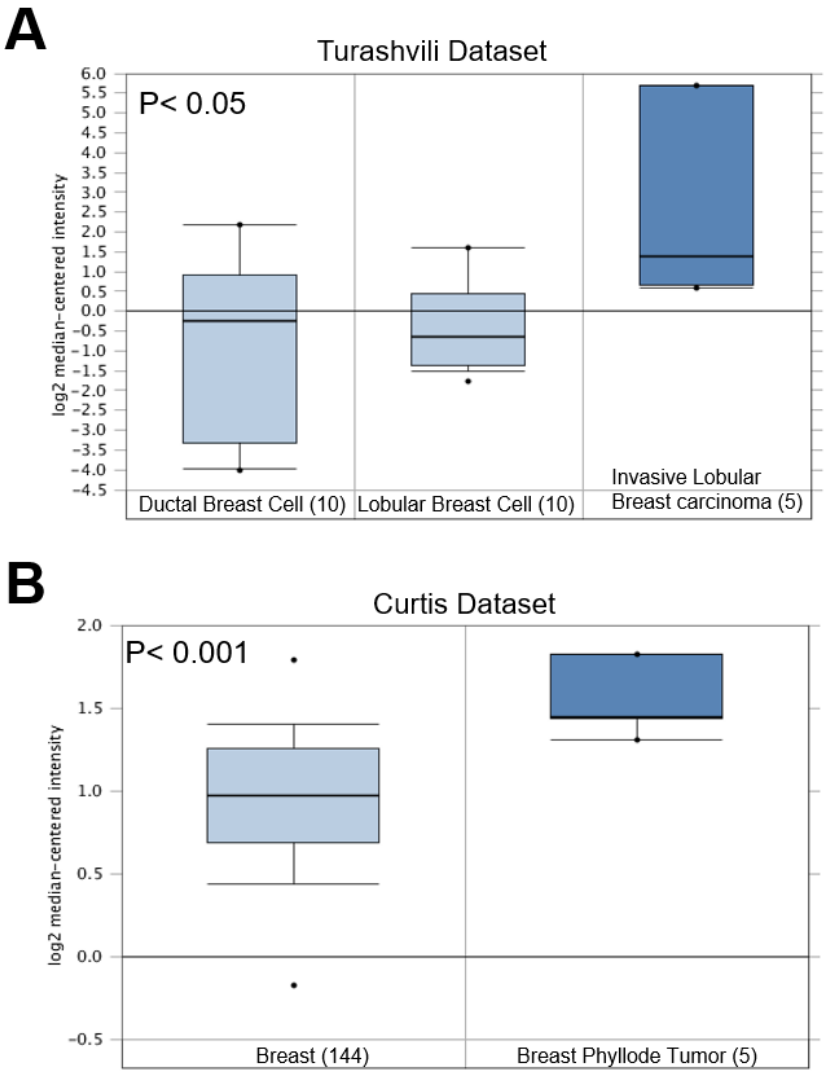
Experiments were repeated at least twice. Data are presented as mean ± SD. A Student's t-test (two tailed) was used to compare two groups. $p < 0.05$ was considered statistically significant.

CHAPTER 4. RESULTS

4.1 RBMS3 is correlated with malignance of breast cancer

To determine whether *rbms3* mRNA expression is altered in breast cancer, we interrogated gene expression datasets of human breast cancer samples from Oncomine database. According to Turashvili and colleagues' dataset, high level of *rbms3* was found in the invasive lobular breast carcinoma compared to ductal and lobular breast cells (Figure 4.1A). In Curtis dataset, 5 phyllodes tumor samples and 144 paired normal breast samples were analyzed (Figure 4.1B). The mRNA expression of *rbms3* was significantly increased in phyllodes tumor samples when compared with normal controls. Interestingly, additional analysis on Finak dataset of 53 breast tumor stroma samples and 6 normal breast stroma samples revealed that the mRNA level of *rbms3* was found significantly higher in the invasive breast carcinoma stroma compared with normal breast stroma (Figure 4.1C). In addition, *rbms3* transcript was found to be to be absent in noninvasive luminal breast cancer cell lines MCF7, T47D and SK-BR-3 (Figure 4.2). Our microarray dataset indicated significant upregulation of *rbms3* in the triple-negative breast cancer (TNBC) cells compared with luminal BC MCF7 cells (Figure 4.3A). Based on gene expression profiles from Cancer Cell Line Encyclopedia (CCLE) dataset, *rbms3* was also markedly increased in the BLBC cell lines (MDA-MB-157, MDA-MB-231, Hs578T and BT549) compared to luminal breast cancer cell lines (MCF7, T47D and ZR-751) (Figure 4.3B). IHC staining from the Human Protein Atlas database showed higher protein expression of RBMS3 in myoepithelial cells as compared to ajacent luminal cells (Figure 4.4). Using UCSC Xena online tool, we found that high expression of *rbms3* predicts lower overall survival in 1273 samples from TCGA BRCA dataset (Figure 4.5A). Furthermore, Kaplan-Meier Plotter

(KM Plotter) revealed that higher expression of *rbms3* correlated with lower distant metastasis free survival (DMFS) using microarray data from 1746 breast cancer samples (Figure 4.5B). KM Plotter also predicted lower DMFS for low expression of *rbms3* in 458 samples of grade 3 breast cancer patients (Figure 4.5C). Altogether, these data suggest that RBMS3 correlates with breast cancer malignancy.



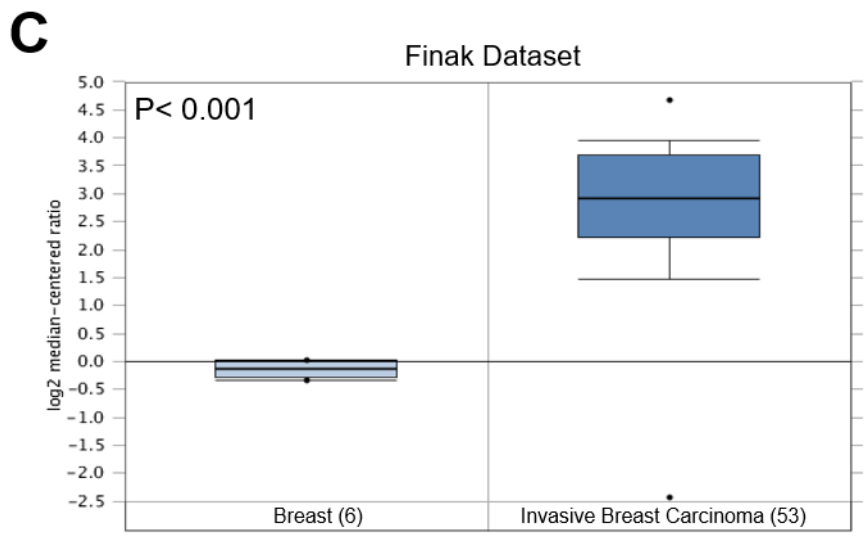


Figure 4.1 *Rbms3* is upregulated in invasive breast carcinoma

(A) Comparison of *rbms3* mRNA expression between invasive breast tissue and normal breast cells in Turashvili dataset (P<0.05). (B) Comparison of *rbms3* mRNA expression between normal breast cells and breast phyllodes tumor in Curtis dataset (P<0.001). (C) Comparison of *rbms3* mRNA expression between the stroma of invasive breast carcinoma versus that of normal breast tissue in Finak dataset (P<0.05). Two-sample t-test was used to compare rbms3 gene expression between carcinoma and normal samples. The expression values were log2-transformed median-centered ratio. (data from Oncomine database)

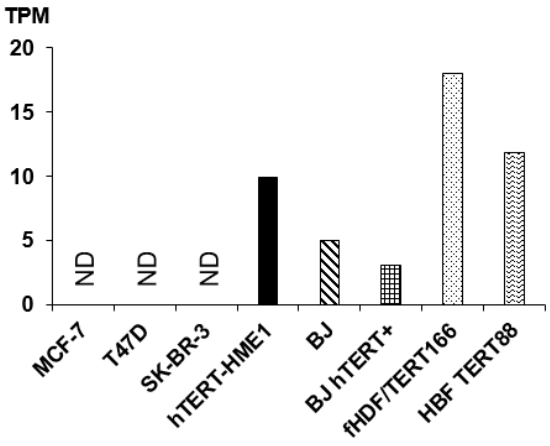


Figure 4.2 *Rbms3* is undetected in luminal and HER2$^+$ subtype BC cell lines
Rbms3 mRNA in luminal MCF7 and T47D breast cancer cells, HER2-overexpressing SK-BR-3 breast cancer cells, hTERT-HME1 normal breast cells and four fibroblast cell lines (BJ, BJ hTERT+, fHDF/TERT166 and HBF TERT88) in the Human Protein Atlas database. Expression quantification was shown as TPM (Transcripts Per Million). ND indicates no transcripts detected.

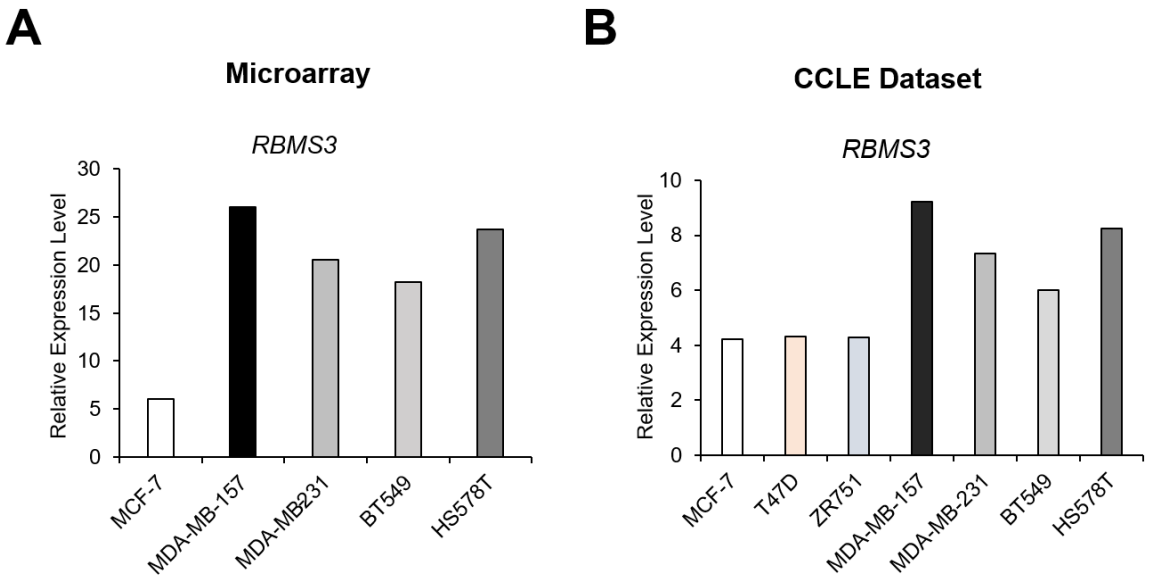


Figure 4.3 *Rbms3* is significantly increased in metastatic TNBC cell lines
(A) Analysis from microarray expression dataset contained 1 luminal (MCF7) and 4 triple-negative/basal-like (MDA-MB-157, MDA-MB-231, HS578T and BT549) breast cancer cell lines.

(B) Analysis from CCLE expression dataset contained 3 luminal (MCF7, T47D and ZR-751) and 4 triple-negative/basal-like (MDA-MB-157, MDA-MB-231, HS578T and BT549) breast cancer cell lines.

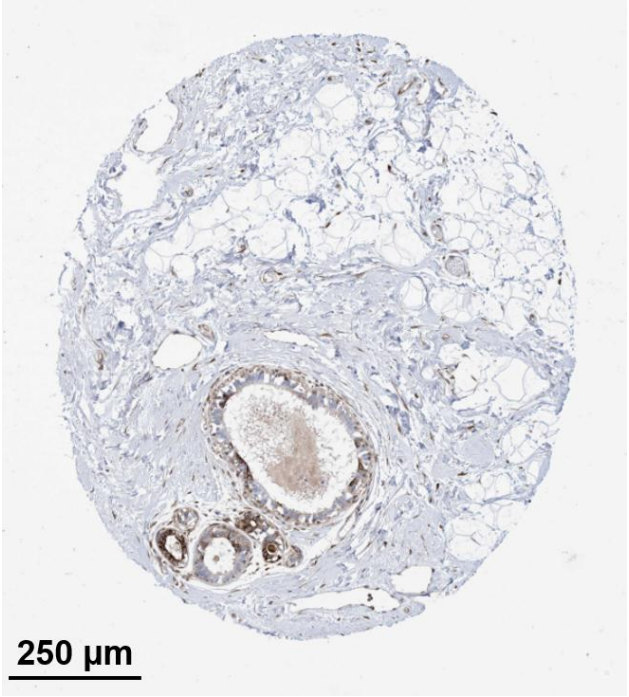
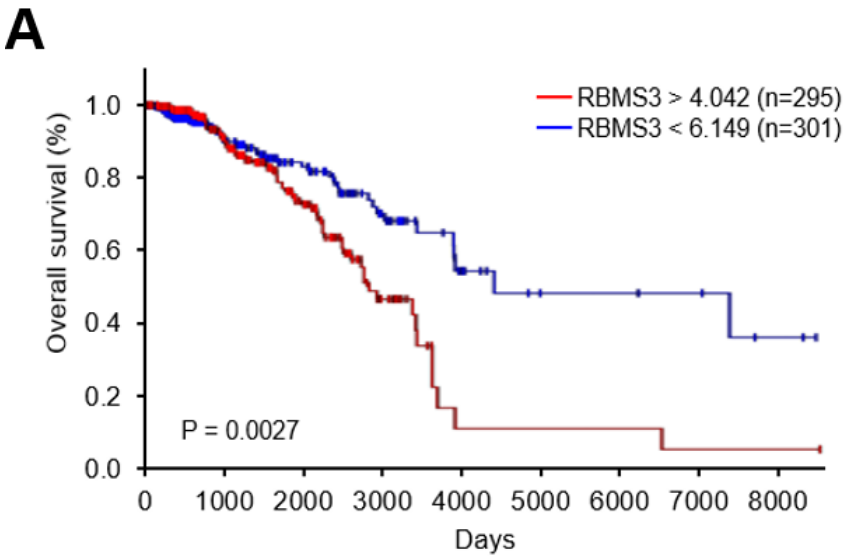


Figure 4.4 RBMS3 is found in the basal-like myoepithelial cells but not in adjacent luminal cells

Representative IHC stain of RBMS3 in normal breast tissue from the Human Protein Atlas database. Brown color stains for RBMS3 protein. Scale bar, 250 µm.



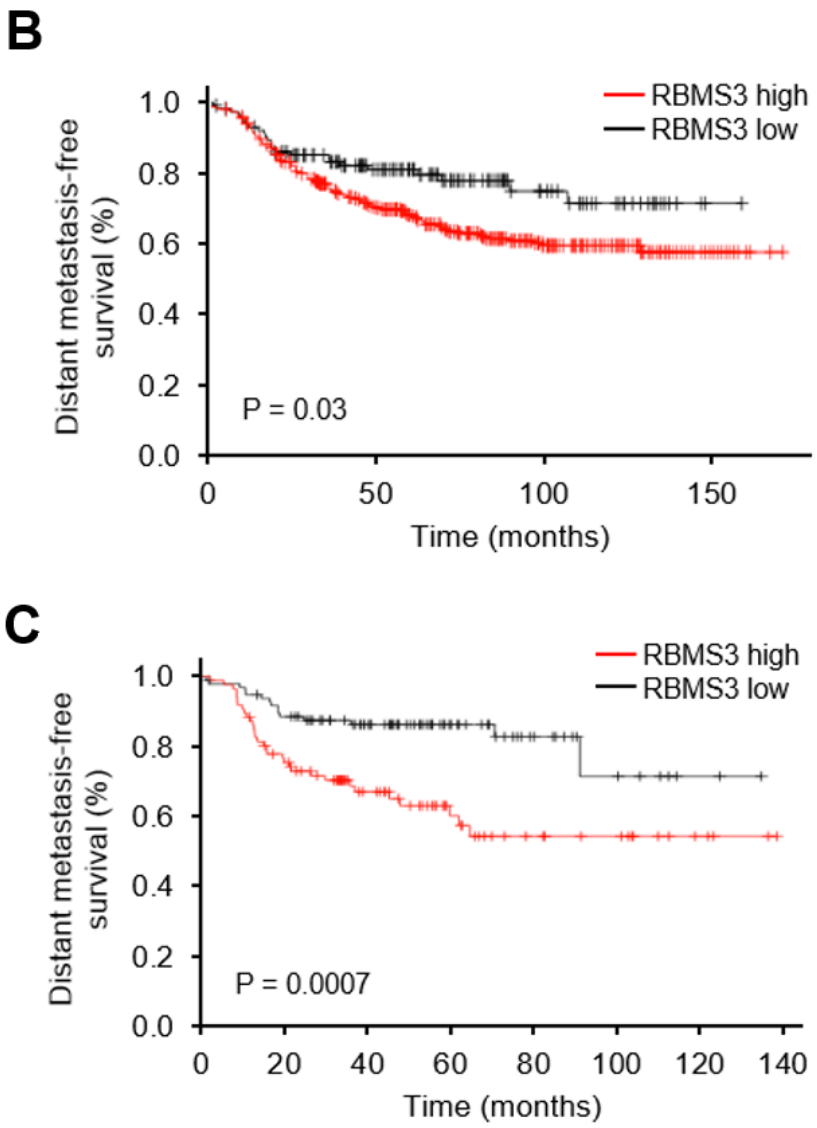


Figure 4.5 *Rbms3* is correlated with poor survival in BC patients
(A) Kaplan Meier plot of overall survival based on *rbms3* gene expression in 1247 samples from TCGA breast cancer (BRCA) dataset using UCSC Xena tool. Results were shown in quartiles. p value < 0.05

(B)Kaplan Meier plot of distant metastasis-free survival based on *rbms3* gene expression in 1746 breast cancer patients. p value <0.05 [145]

(C) Kaplan Meier plot of distant metastasis-free survival based on *rbms3* gene expression in 458 breast cancer patients (grade 3). p value < 0.001 [145]

4.2 RBMS3 induces cell migration, invasion, EMT and CSC traits

EMT program can impart several traits that are essential to the malignant progression of carcinoma cells, including tumor-initiating properties, motility, the ability to disseminate and increased tolerance to conventional chemotherapeutics [29, 146-148]. To investigate the functional effects of RBMS3, ectopic expression of RBMS3 was induced in two luminal breast tumor cell lines, T47D and MCF7, which contain little or no endogenous RBMS3. To study whether RBMS3 affects breast cancer cell viability, we performed cell viability assay in MCF7 and T47D cells stably expressed control or RBMS3 vector. MCF7-RBMS3 and T47D-RBMS3 cells demonstrated minimal effect on cell growth compared to control cells over the 96-hour interval examined (Figure 4.6A and B). Next, we tested migration and invasiveness of these cells using Boyden chamber. RBMS3 expression markedly increased MCF7 cell migration and invasive capacity (Figure 4.7). Similar results were also observed in T47D and MCF10A cells (data not shown). RBMS3 induced morphologic changes reminiscent of EMT, including downregulation of epithelial marker *CDH1* and upregulation of mesenchymal markers (*TWIST1*, *TWIST2*, *SNAI1*, *SNAI2*, *FOXC1* and *FN1*) in T47D cells. In addition, T47D-RBMS3 cells lost luminal markers *CD24* and gained expression of stem cell molecule *Sox2* (Figure 4.8A). Similar qPCR results were also seen in MCF7-RBMS3 cells (Figure 4.8B). Consistently, RBMS3 expression decreased the protein levels of epithelial marker E-cadherin, and induced the protein levels of stem cell molecules (Sox2 and LSD1) in MCF7 and T47D cells (Figure 4.9). These data suggest RBMS3 is critical in controlling the expression of EMT and pluripotency genes.

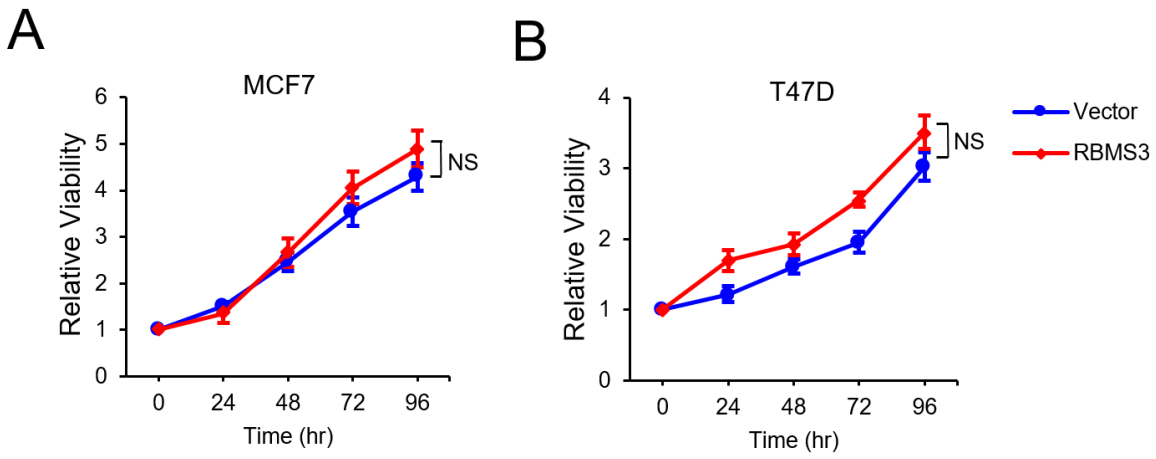


Figure 4.6 RBMS3 has minimal effect on MCF7 or T47D cell growth
Graphic representation of cell growth rates by MCF7 (A) and T47D (B) cells stably expressed RBMS3 or control vector. Cell viability was measured daily over a 4-day period. Presented data are the mean ± SD from three independent experiments.

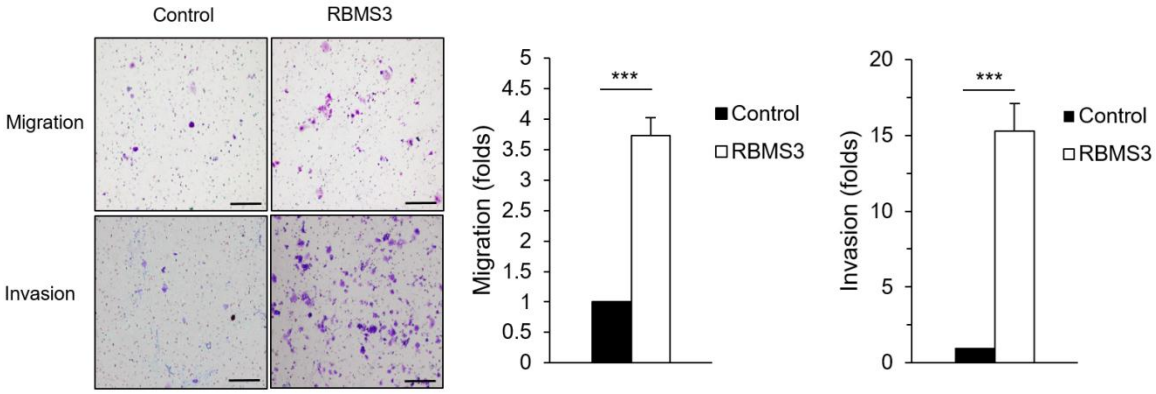


Figure 4.7 RBMB3 promotes MCF7 cell migration and invasion
Graphic representation of the invasiveness of T47D cells stably expressed control or RBMS3 vector using a modified Boyden Chamber migration (A) or invasion (B) assay. Quantification is shown in the right panel. Presented data are the mean ± SD from two independent experiments in triplicates, with *** indicates p value < 0.001 when comparing with control values. Scale bars, 100 μm.

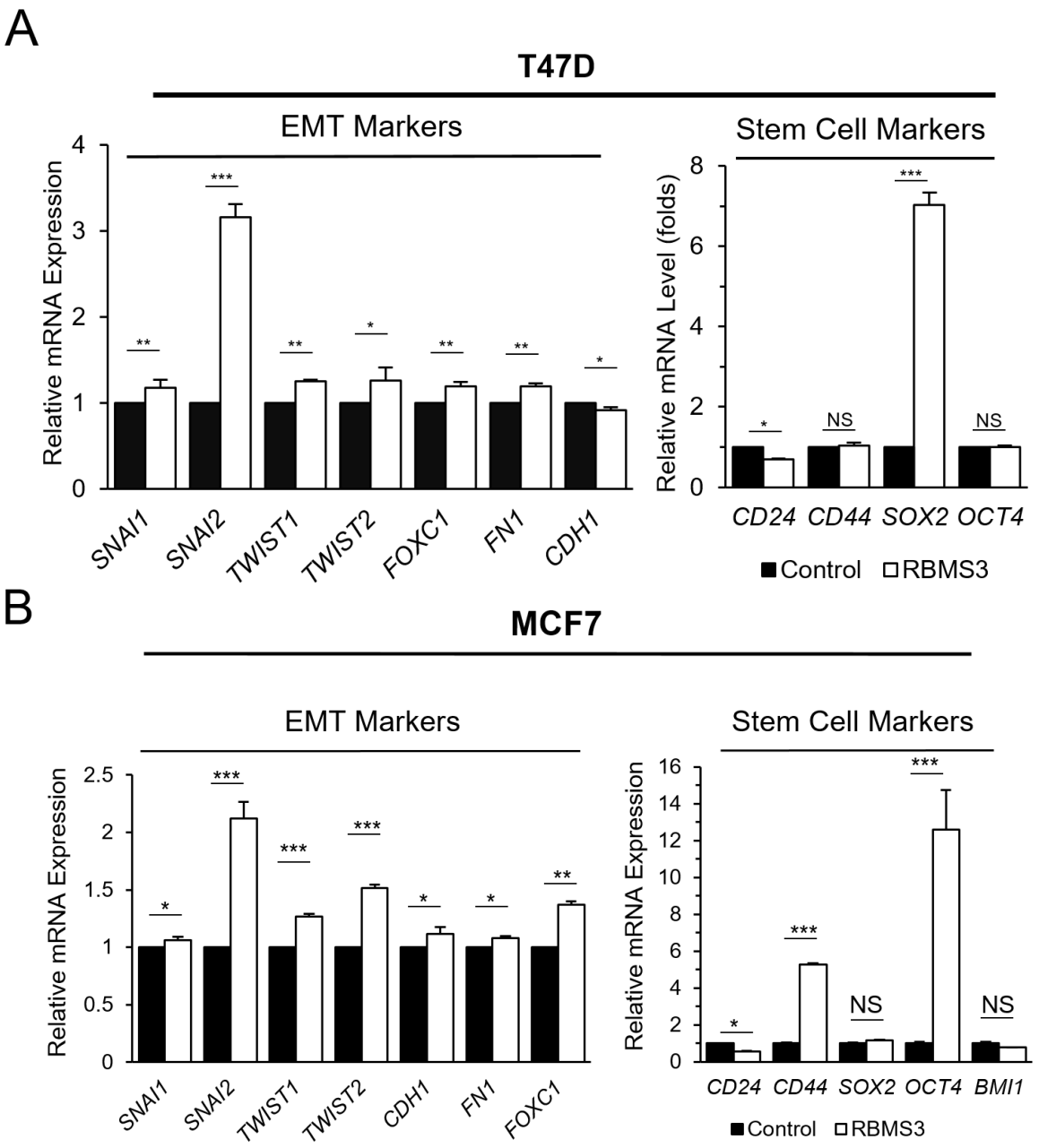


Figure 4.8 RBMS3 induces the mRNA levels of EMT and stem cell markers
(A-B) Real-time PCR analysis of the mRNA level of several EMT markers and stem cell markers in T47D cells or MCF7 cells expressed control or RBMS3 vector. Data are shown as mean ± SD from three independent experiments in triplicates.

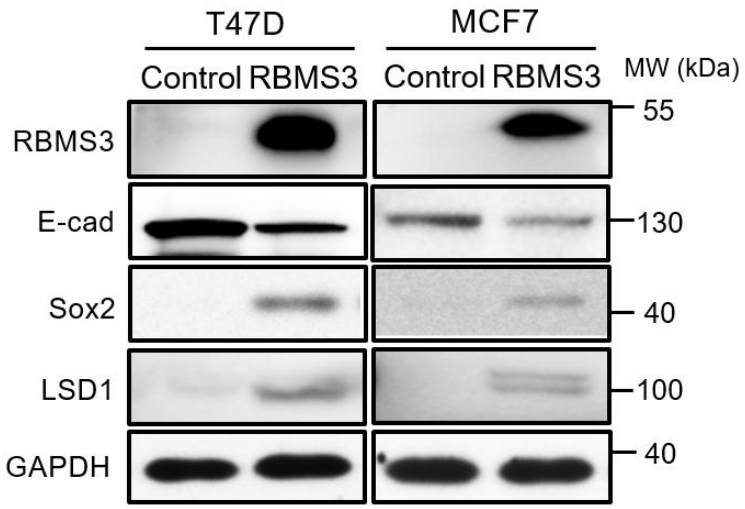


Figure 4.9 RBMS3 induces protein expression of EMT and stem cell markers

Western blot analysis of EMT marker (E-cadherin) and stem cell markers (Sox2 and LSD1) in T47D and MCF7 cells stably expressed control or RBMS3 vector. GAPDH was loaded for normalization.

4.3 RBMS3 promotes cancer stem cell (CSC)-like characteristics

Cancer cells that develop the ability to undergo EMT lose anchorage dependence and thus can detach from the primary tumor. Soft-agar assay was performed in T47D cells to determine whether RBMS3 affects anchorage-dependent growth of these cells. RBMS3 expression resulted in ~ two-fold increase and ~ five-fold increase in the number of soft-agar colonies in T47D and MCF7 cells, respectively (Figure 4.10). Tumorsphere formation is largely dependent upon the self-renewal and tumorigenic abilities of stem/progenitor cells, known as cancer stem cells (CSCs), to survive and growth in serum-free suspension. Cells are grown in serum-free, non-adherent conditions in order to CSC/progenitor cells since only CSC/progenitor cells can survive and proliferate under such environment . CSCs is known as the main reason for cancer recurrence, metastasis and therpeutic resistance [149]. To study the role of RBMS3 in tumorsphere formation, we cultured MCF7 cells expressed control or RBMS3 vector in serum-free suspension medium for two weeks and then examined tumorsphere formation of these cells by looking at their size and density.

MCF7 control cells formed moderate numbers of mammospheres with similar average sizes (30-100 μm). Surprisingly, MCF7-RBMS3 cells appeared to form mammospheres of varying sizes. Quantification of tumorsphere counts suggested that while RBMS3 expression led to moderate decrease in mammospheres of sizes above 30 μm, it also led to dramatic increase in small irregular cell aggregates/clusters below 30 μm (not strictly consider as tumorsphere) or even mini cell clusters (Figure 4.11). Formation of varying sizes of mammospheres, especially those mini clusters, is particularly favorable for tumor cell dissemination in the circulation. Similar results were also observed in T47D-RBMS3 cells (data not shown). These results, together with qPCR and western blot results, indicate that RBMS3 may enhance CSC-like properties in luminal breast cancer cells.

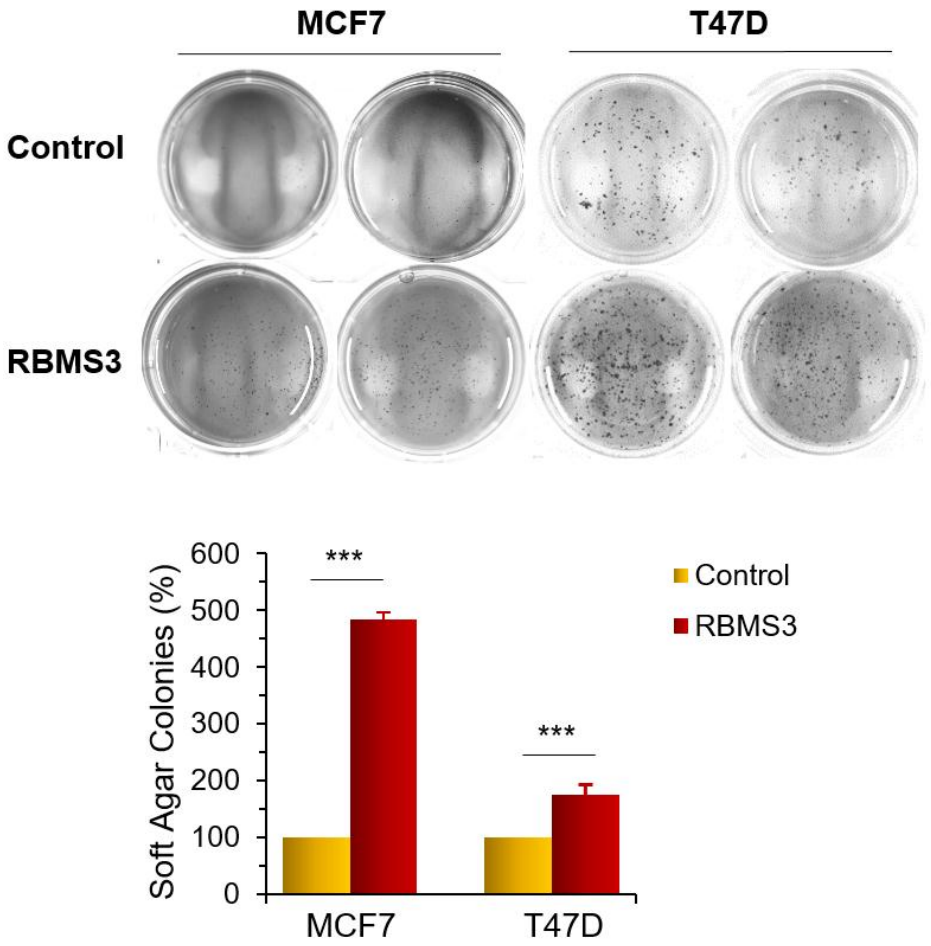


Figure 4.10 RBMS3 promotes colony formation in MCF7 and T47D cells

Data of colony formation assay are presented as a percentage of vector control cell lines. Data are shown as mean ± SD in two independent experiments in triplicates. (***$p < 0.001$)

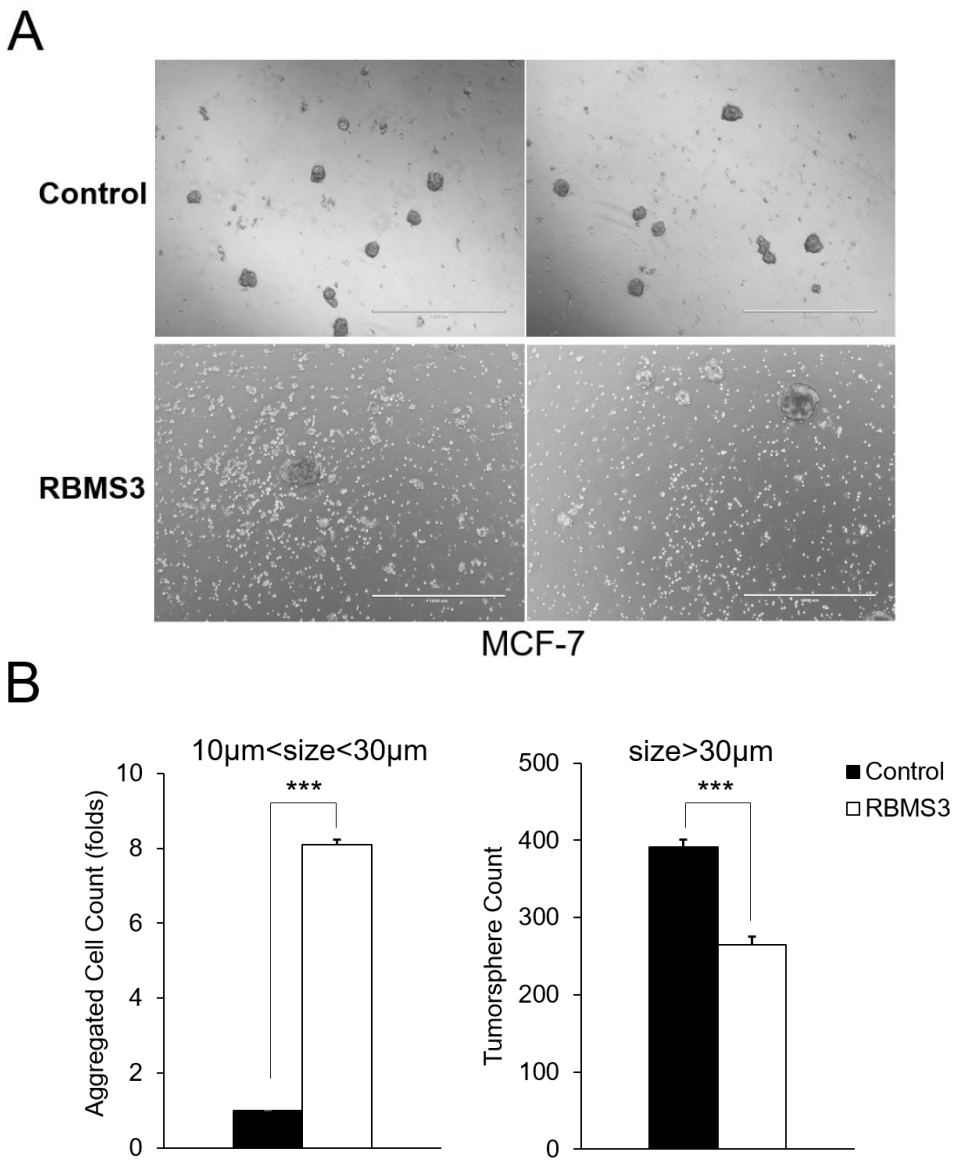


Figure 4.11 RBMS3 promotes tumorsphere formation in BC cells

(A) Tumorsphere formation was assessed in MCF7 cells stably expressed control or RBMS3 vector. Representative images of tumorspheres are shown. Scale bars, 1000 μm. (B) Data of tumorsphere formation assay are presented as counts (size over 30 μm) or folds of control vector values (size between 10-30 μm), with mean ± SD of two independent experiments performed in triplicates (***$p < 0.001$).

4.4 Loss of RBMS3 Suppresses EMT and CSC population

To further explore the function of RBMS3 in TNBC cells, we established stable RBMS3-knockout clones in BT-549, Hs578T, MDA-MB-157 and MDA-MB-231 cells by CRISPR/Cas9 technique. BT-549 cells appeared with stellate projections, whereas BT549-RBMS3 KO cells exhibited a marked change in morphology, with rounded/polygonal shape when cultured at low density (Figure 4.12). To investigate whether RBMS3 affects breast cancer cell growth, we measured cell growth rates in BT-549 cells by cell viability assay. Over the time course examined, RBMS3 knockout demonstrated significant decrease on cell growth in BT-549 cells (Figure 4.13). We also measured cell growth in Hs578T cells by cell counting. Knockout of RBMS3 resulted in significant decrease in growth rate in Hs578T cells (Figure 4.14). To investigate the migratory ability mediated by RBMS3 in TNBC cells, transwell assay was performed BT-549 and MDA-MB-157 cells. Indeed, loss of RBMS3 significantly inhibited cell migration (Figure 4.15). This finding was supported by in vitro wound healing assay in MDA-MB-231 cells. Closure of the scratch wound required significantly longer time in RBMS3-depleted cells than in control cells. Statistical analysis indicated that migratory activity of RBMS3-depleted cells was 60% lower than that of control cells (Figure 4.16). Loss of RBMS3 led to significant upregulation of epithelial markers (*CDH1* and *FOXA1*) and downregulation of several mesenchymal markers (*SNAI2*, *TWIST1*, *TWIST2*, *CDH2*, *FN1* and *FOXC1*) in BT549 cells (Figure 4.17). Importantly, RBMS3 depletion reduced the protein levels of epithelial marker E-cadherin, while increased the protein levels of mesenchymal markers (Slug, Twist, Vimentin, Snail and N-cadherin) in BT549 and MDA-MB-157 cells, as confirmed by western blotting (Figure 4.18). To test whether RBMS3 affected CSC characteristics, BT549 control and RBMS3 knockout cells were subjected to FACS analysis using $CD44^{high}/CD24^{low}$ and $CD49f^{high}/CD24^{low}/EpCAM^{+}$ as surface markers. The $CD44^{high}/CD24^{low}$ and $CD49f^{high}/CD24^{low}/EpCAM^{+}$ CSC populations (represented by the

upper left quadrant) were both dramatically decreased in RBMS3-depleted cells vs. control cells, confirming that RBMS3 positively regulates CSC traits (Figure 4.19). Altogether, the above findings further support the role of RBMS3 as a crucial factor in promoting breast cancer cell migration, invasion, EMT as well as CSC characteristics.

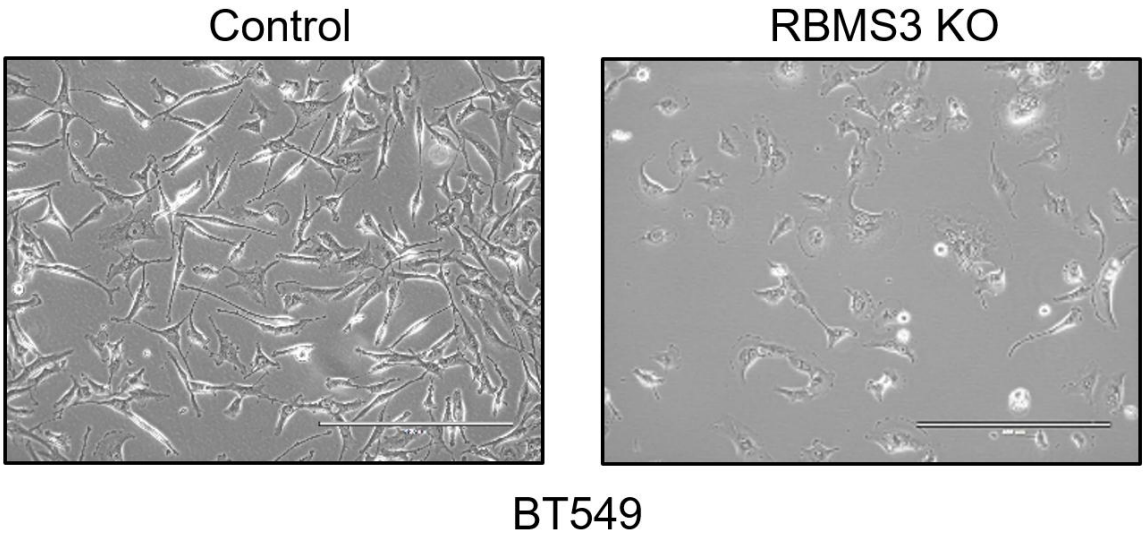


Figure 4.12 Loss of RBMS3 reduces mesenchymal phenotype in BT549 cells
Representative images of BT549 cells with or without loss of RBMS3. At low cell density, BT549 cells exhibit elongated spindle-like shape, whereas RBMS3 KO cells show shortened, polygonal and spread-out shape. Scale bars, 400 μm.

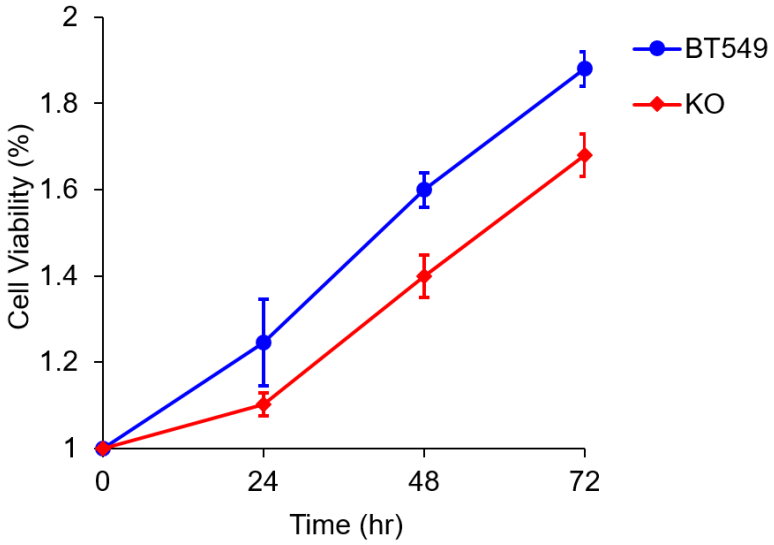


Figure 4.13 Loss of RBMS3 significantly inhibits BT549 cell viability
Graphic representation of cell viability in BT549 cells with or without loss of RBMS3 over a 3-day period measured by CellTiter-Glo 2.0 Assay. Presented data are the mean ± SD from three independent experiments. (*$p < 0.05$)

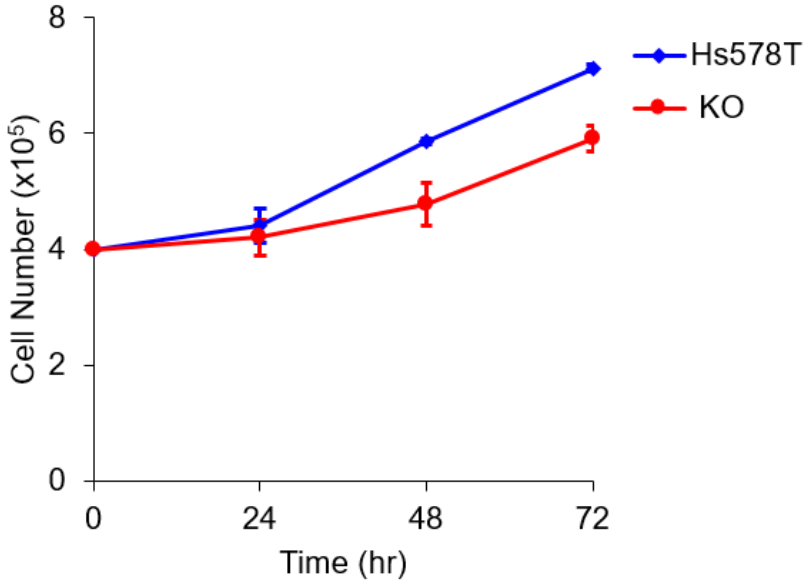


Figure 4.14 Loss of RBMS3 significantly decreases Hs578T cell growth
Graphic representation of cell growth rates in Hs578T cells with or without loss of RBMS3 over a 3-day period measured by cell counting. Presented data are the mean ± SD from three independent experiments. (*$p < 0.05$)

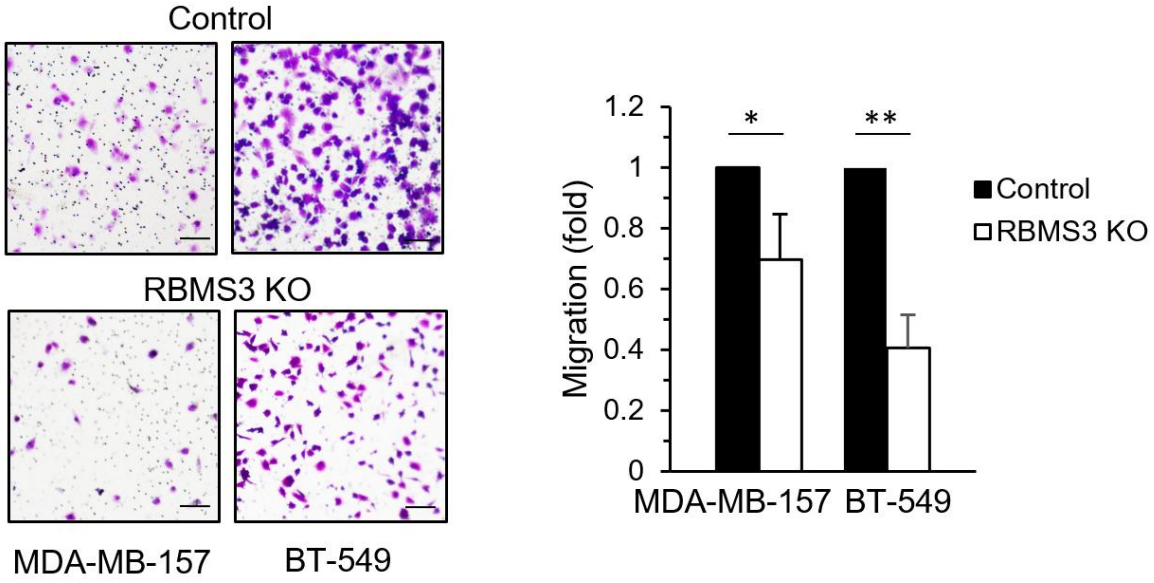


Figure 4.15 Loss of RBMS3 inhibits BT-549 and MDA-MB-157 cell migration
Graphic representation of the invasiveness of MDA-MB-157 and BT-549 cells with or without loss of RBMS3 using a modified Boyden Chamber migration assay. Quantification is shown in the right panel. Presented data are the mean ± SD from two independent experiments in triplicates, with ** indicates p value < 0.01, *p <0.05 when comparing with control values. Scale bars, 100 μm.

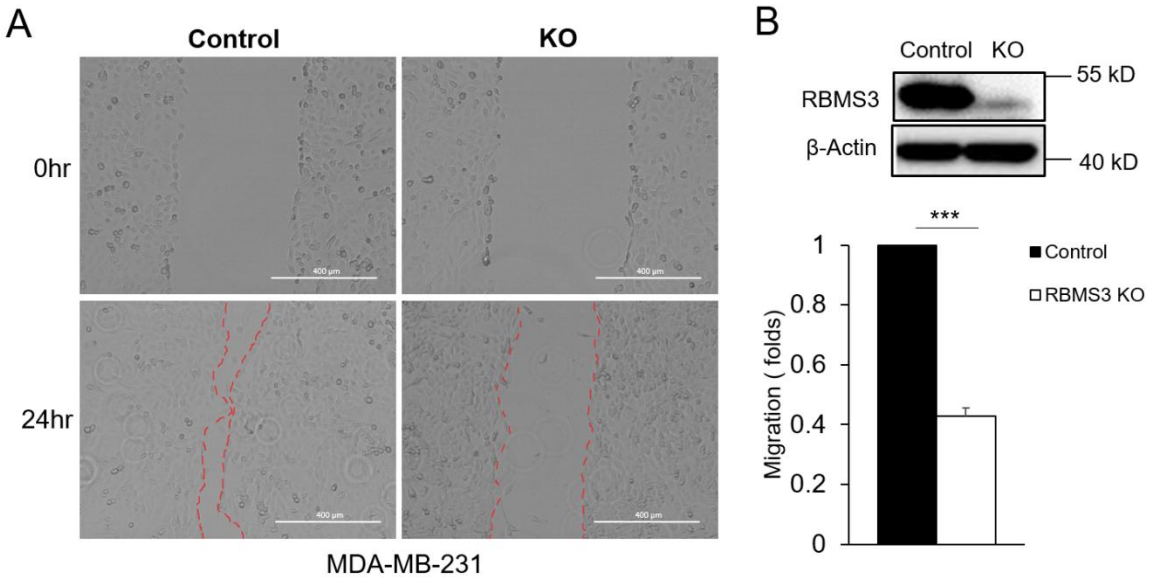


Figure 4.16 Loss of RBMS3 prevents wound closure in MDA-MB-231 cells

Representative images of scratch assay in MDA-MB-231 cells with or without loss of RBMS3. Distance of wound closure was measured at 0 and 24h. Red dashed lines show the margins of migrating cells. Scale bars, 1000 μm. (G) Western blot analysis of RBMS3 knockout in MDA-MB-231 cells and statistical analysis of 10 random measurements of scratch assay. (***$p < 0.001$) (H) Migratory ability of MDA-MB-231 and BT549 cells and the corresponding RBMS3 knockout cells were analyzed by transwell migration assay. Scale bars, 200 μm. (I) Invasive ability of MDA-MB-231 and BT549 cells and the corresponding RBMS3 knockout cells were analyzed by transwell invasion assay. Scale bars, 200 μm.

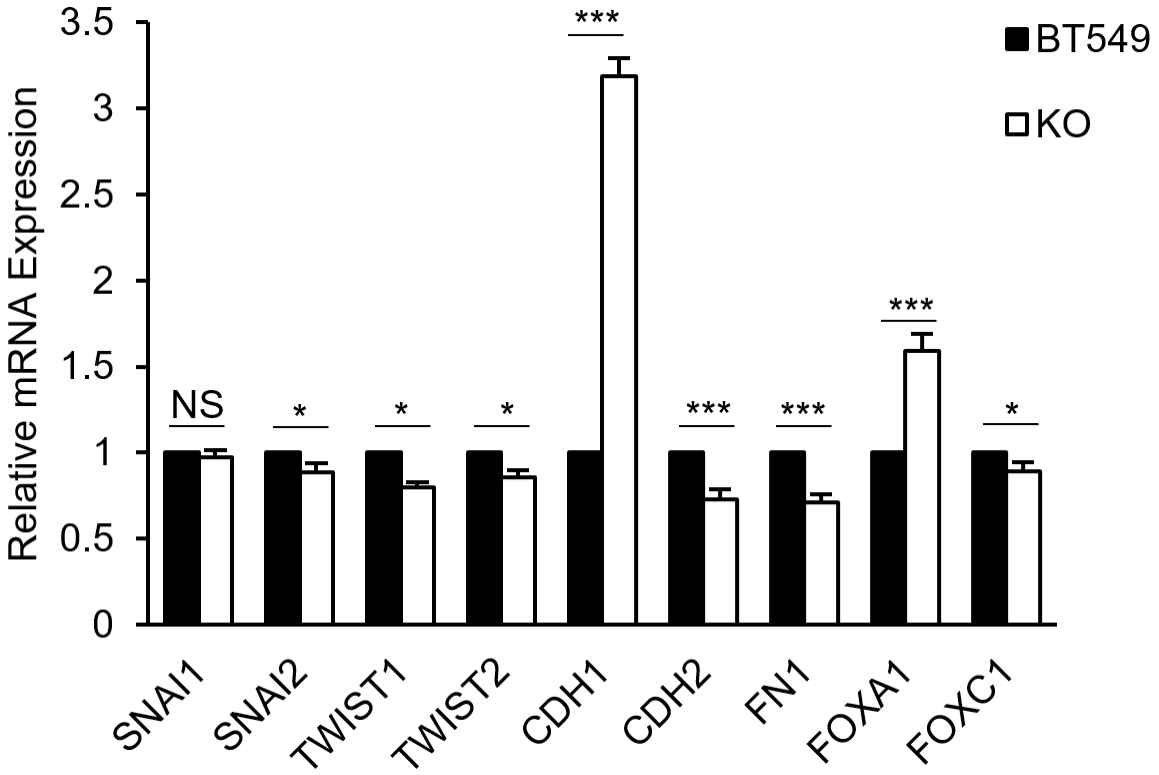


Figure 4.17 Loss of RBMS3 reduces mesenchymal markers and increases epithelial markers at mRNA level

Real-time PCR analysis of the mRNA level of several mesenchymal markers (*SNAI1, SNAI2, TWIST1, TWIST2, CDH2, FN1, FOXC1*) and epithelial markers (*CDH1, FOXA1*) in BT549 cells with or without loss of RBMS3. Data are shown as mean ± SD from two independent experiments in triplicates. (*p < 0.05, **p < 0.01, *** < 0.001, NS = not significant)

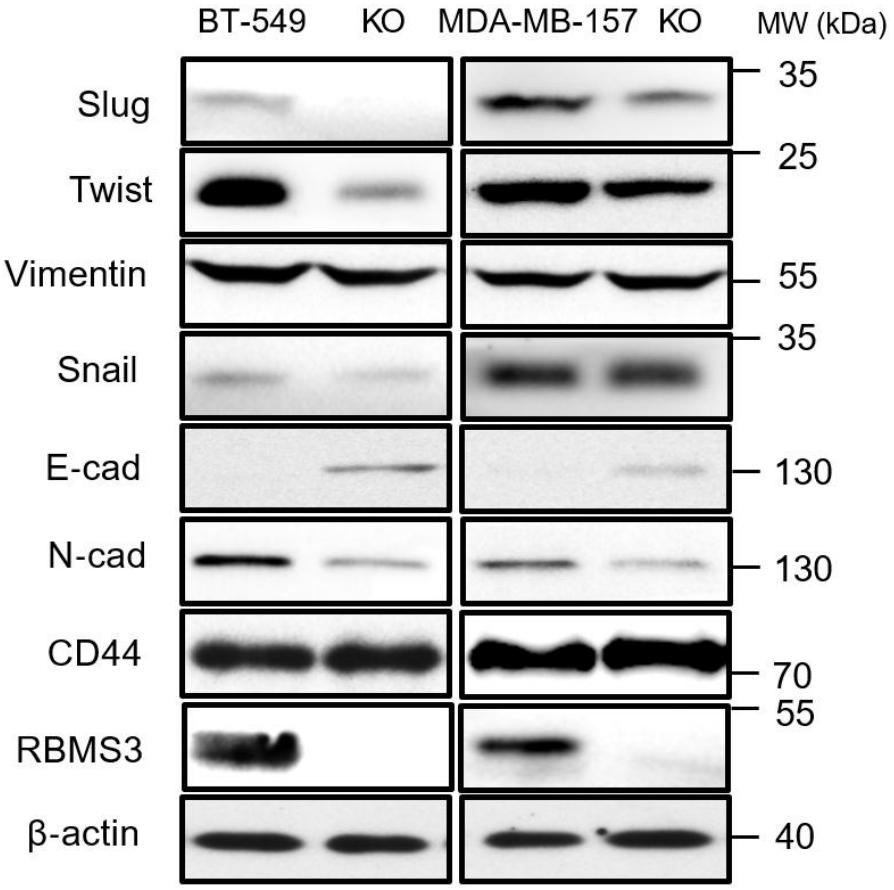


Figure 4.18 Loss of RBMS3 reduces mesenchymal markers and increases epithelial markers at protein level

Western blot analysis of EMT markers in BT549 and MDA-MB-157 cells with or without loss of RBMS3. β-actin was loaded as control.

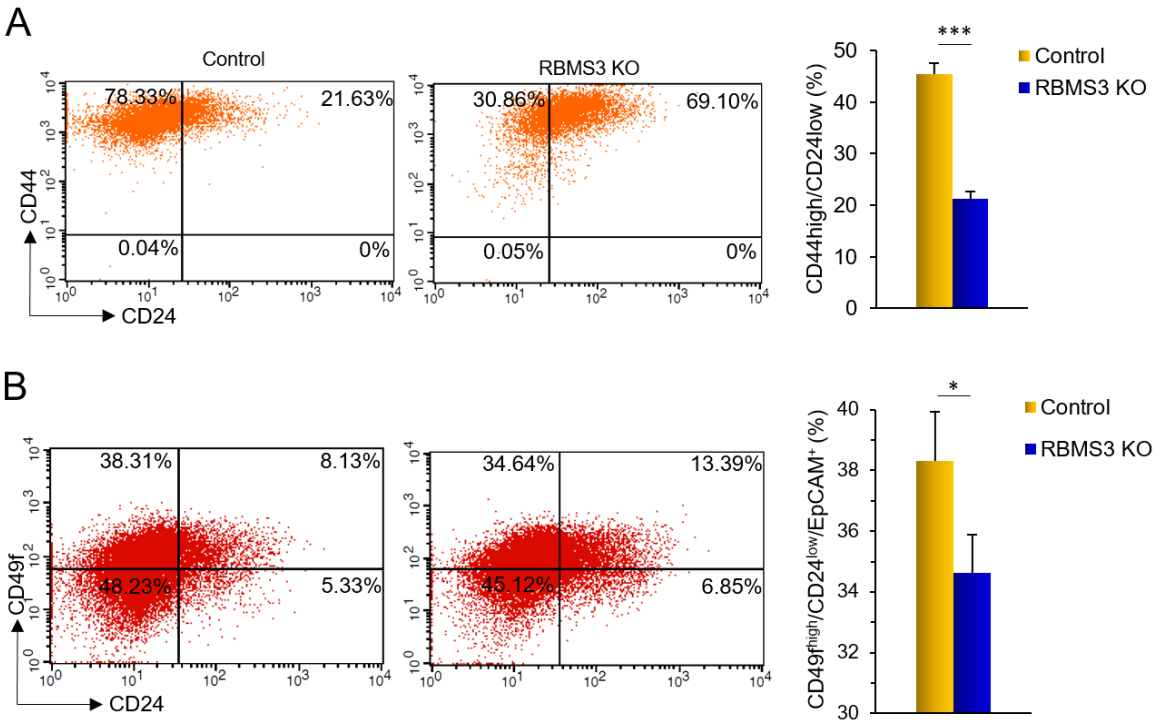


Figure 4.19 Loss of RBMS3 reduces $CD44^{high}/CD24^{low}$ and $CD49f^{high}/CD24^{low}/EpCAM^{+}$ CSC populations

Representative FACS images from BT549 cells with or without loss of RBMS3 using $CD44^{high}/CD24^{low}$ or $CD49f^{high}/CD24^{low}/EpCAM^{+}$ as surface markers. Quantification of CSC populations are shown in the right panel; Presented data are the mean ± SD from three experiments. (***$p < 0.001$)

4.5 RBMS3 is correlated with TGF-β Signaling

To further investigate the molecular mechanism of RBMS3 involved in the metastasis of breast cancer, T47D cells expressed control or RBMS3 vector were selected for RNA-seq analysis to screen for potential targets of metastasis. GO biological analysis of deregulated genes (DEGs) between RBMS3 vs. control list revealed top 20 significant pathways, including genes encoding ECM-associated proteins, GPCR ligand binding and chemotaxis-associated proteins, etc. (Figure 4.20). A subset of enriched terms was selected for network plot (Figure 4.21) and protein-protein interaction networks were shown using Metascape platform (Figure 4.22-23 and Table 2). Table 3 shows a list of top 20 DEGs discovered in the RNA-seq dataset. Gene set enrichment analysis (GSEA) indicated an

enrichment of pathway of EMT, inflammatory response, hypoxia and TGF-β signaling in T47D-RBMS3 cells compared with T47D control cells (Figure 4.24). In addition, positive correlation between RBMS3 and apical junction pathway as well as other pathways was also found in T47D-RBMS3 cells (Table 4). Significant correlation of RBMS3 with EMT and TGF-β signatures was also observed in the BRCA dataset from TCGA database using GEPIA2 online analysis tool (Figure 4.25). A8301 prevents phosphorylation of Smad2/3, leading to decreased Smad-dependent TGF-β signaling. TGF-β1 treatment inhibited the expression of its downstream target *CTGF* in T47D cells, which was potentiated by Alk5 inhibitor A8301, suggesting that TGF-β1-mediated downregulation of *CTGF* is mediated mainly through non-Smad pathways. Overexpression of RBMS3 in T47D cells led to significantly downregulated *CTGF* expression, which was enhanced by TGF-β1 treatment. A8301 treatment, however, failed to reverse downregulation of *CTGF*, indicating that RBMS3-mediated *CTGF* downregulation was mainly regulated through non-Smad pathways (Figure 4.26). We also noticed that overexpression RBMS3 facilitated TGF-β1-mediated downregulation of E-cadherin in MCF7 cells (Figure 4.27).

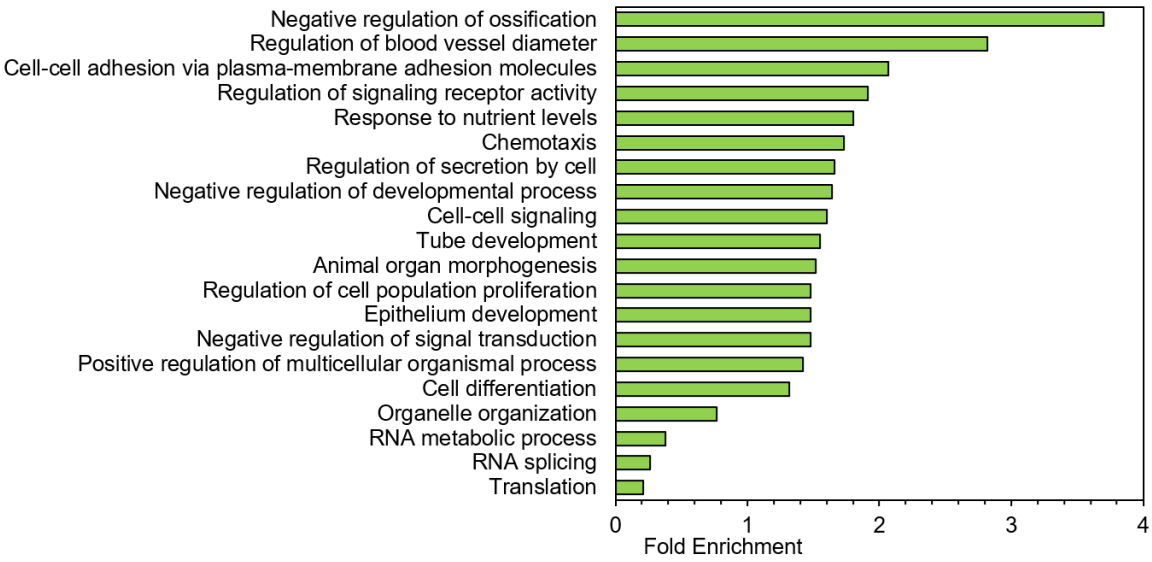


Figure 4.20 Functional enrichment analysis of deregulated gene set (DEGs)
The DEG list of T47D-RBMS3 cells (vs. T47D control) was subjected to enrichment analysis using Gene Ontology (GO) terms for biological processes, cellular components and molecular function categories annotated by Gene Ontology. Data represented as fold enrichment.

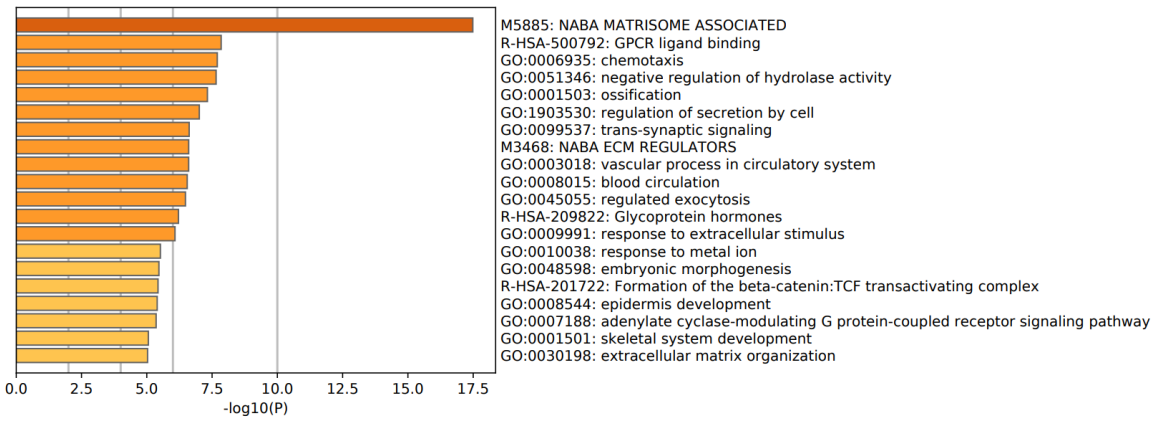


Figure 4.21 Heatmap for GO analysis of DEGs showing top 20 significant pathways
The DEG list of T47D-RBMS3 cells (vs. T47D control) was subjected to pathway and process enrichment analysis with ontology sources: Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway, GO Biological Processes Reactome Gene Sets, Canonical Pathways and CORUM. All genes in the genome were included as the enrichment background. P-value was calculated based on accumulative hypergeometric distribution. Enrichment factor is the ratio between the observed counts and the counts expected by chance. Terms with a p-value < 0.01, a minimum count of 3, and enrichment factor > 1.5 are collected and grouped into clusters based on their similarities. The most statistically significant term within a cluster is chosen to represent the cluster. $p<0.05$ (http://metascape.org/)

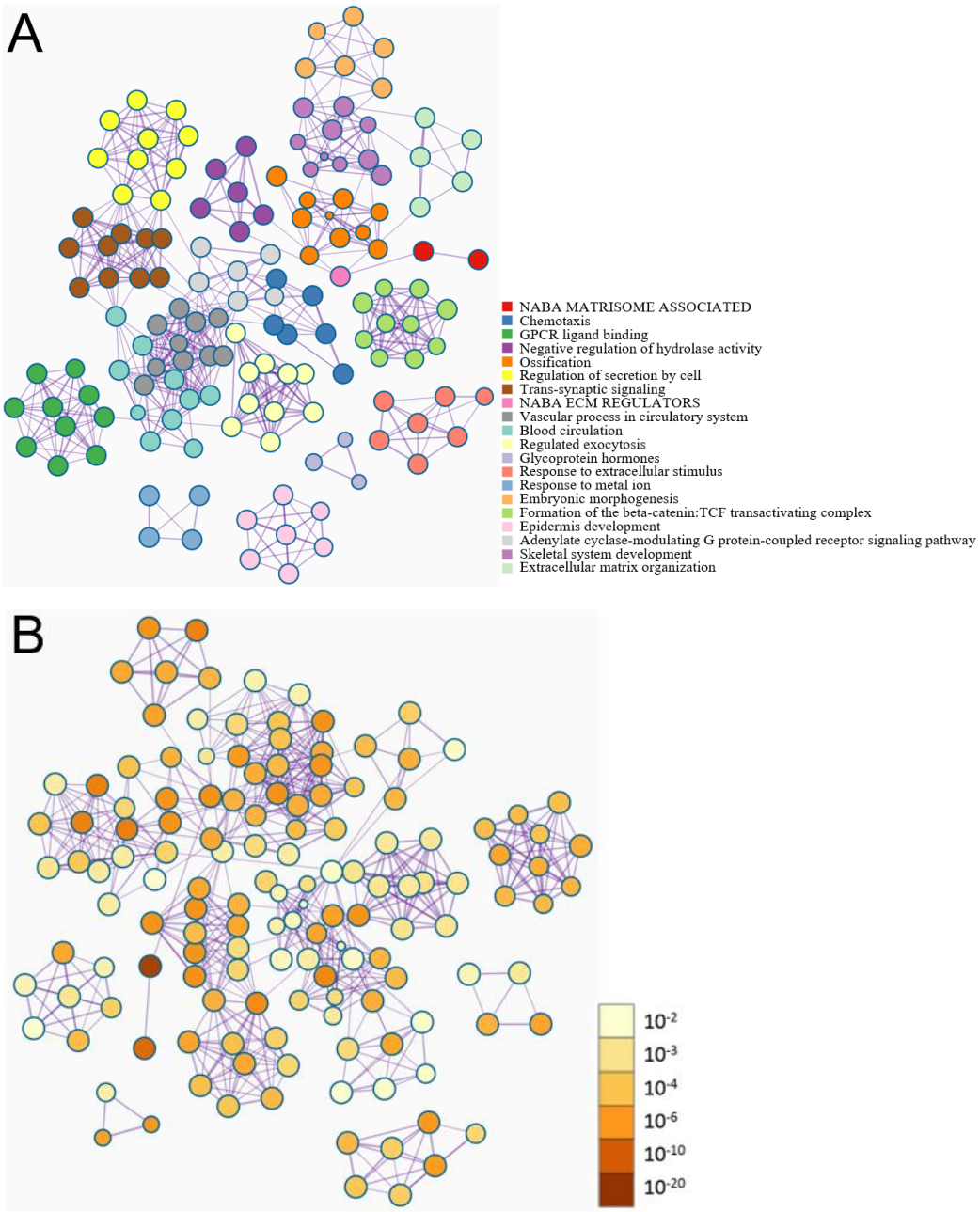


Figure 4.22 Network of Enriched Terms

To further capture the relationships between the terms, a subset of enriched terms was selected as a network plot, where terms with a similarity > 0.3 was connected by edges. The terms with the best p-value from each of the 20 clusters, with constraints that no more than 15 terms per cluster and no more than 250 terms in total, were selected to be visualized. Each node represents an enriched term and is colored by its cluster ID or p-value. a) colored by cluster ID, where nodes that share the same cluster ID are typically close to each other; b) colored by p-value, where terms containing more genes tend to have a more significant p-value. Results displayed in 'cose' layout format. (http://metascape.org/)

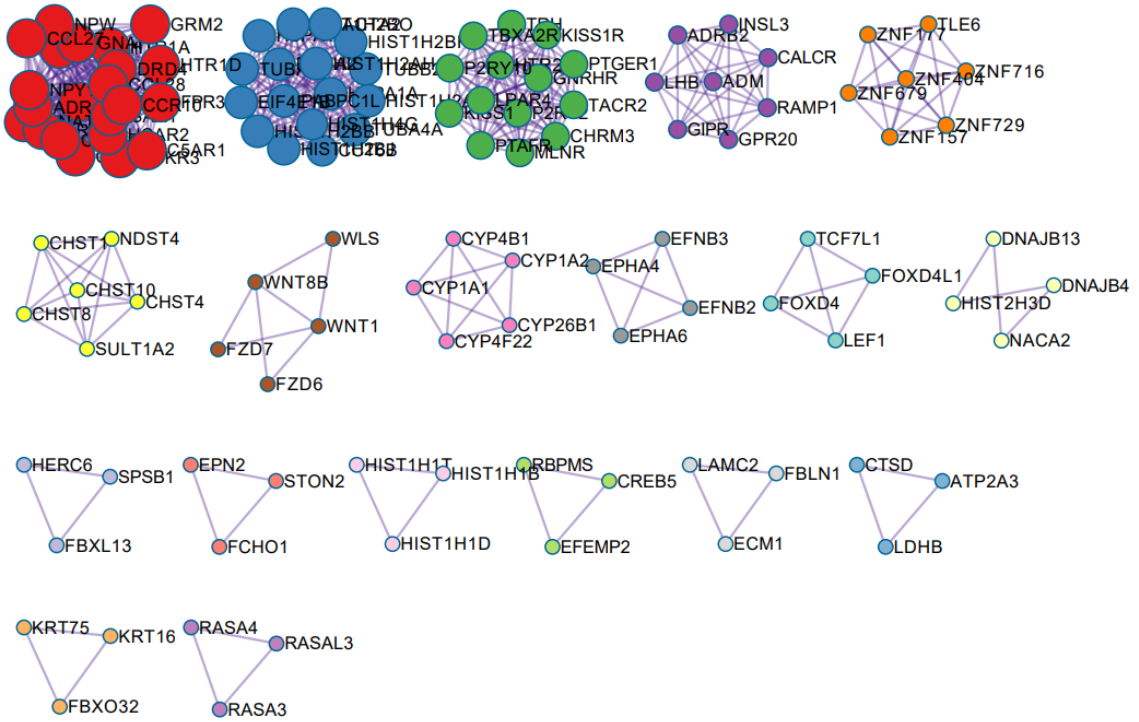


Figure 4.23 Protein-protein Interaction Network and MCODE Components

Protein-protein interaction enrichment analysis was carried out in the DEG list of T47D-RBMS3 cells vs. control cells via the Metascape platform, which utilizes the following databases: BioGrid, InWeb_IM and OmniPath. This network plot contains the subset of proteins that form physical interactions with at least one other member in the list. The Molecular Complex Detection (MCODE) algorithm has been applied to identify and assemble network components for individual gene lists. Pathway and process enrichment analysis was applied to each MCODE component independently, and 2-3 best-scoring terms by p-value are retained as functional description of the corresponding components, shown in Table 3.

Table 2 Protein-protein interaction network and MCODE components identified in the DEG list of T47D-RBMS3 vs. T47D cells

Color	MCODE	GO	Description	Log10(P)
	MCODE_1	R-HSA-418594	G alpha (i) signaling events	-39.3
	MCODE_1	R-HSA-373076	Class A/1 (Rhodopsin-like receptors)	-32.6
	MCODE_1	R-HSA-500792	GPCR ligand binding	-32.3
	MCODE_2	R-HSA-2262752	Cellular responses to stress	-13.3

	MCODE_2	R-HSA-8953897	Cellular responses to external stimuli	-12.4
	MCODE_2	hsa05322	Systemic lupus erythematosus	-11.6
	MCODE_3	R-HSA-416476	G alpha (q) signaling events	-28.8
	MCODE_3	R-HSA-373076	Class A/1 (Rhodopsin-like receptors)	-26.3
	MCODE_3	R-HSA-500792	GPCR ligand binding	-24.2
	MCODE_4	R-HSA-418555	G alpha (s) signaling events	-13.2
	MCODE_4	R-HSA-500792	GPCR ligand binding	-11.2
	MCODE_4	GO:0008277	regulation of G protein-coupled receptor signaling pathway	-9.3
	MCODE_6	GO:0006790	sulfur compound metabolic process	-8.3
	MCODE_6	GO:0005975	carbohydrate metabolic process	-7.2
	MCODE_6	GO:1901137	carbohydrate derivative biosynthetic process	-6.7
	MCODE_7	hsa05217	Basal cell carcinoma	-9.9
	MCODE_7	R-HSA-195721	Signaling by WNT	-9.3
	MCODE_7	GO:0060070	canonical Wnt signaling pathway	-9.3
	MCODE_8	R-HSA-211897	Cytochrome P450 - arranged by substrate type	-12.9
	MCODE_8	R-HSA-211945	Phase I - Functionalization of compounds	-11.8
	MCODE_8	R-HSA-211859	Biological oxidations	-10.2
	MCODE_9	R-HSA-3928665	EPH-ephrin mediated repulsion of cells	-10.8
	MCODE_9	GO:0048013	ephrin receptor signaling pathway	-9.8
	MCODE_9	R-HSA-2682334	EPH-Ephrin signaling	-9.7

	MCODE_12	R-HSA-983168	Antigen processing: Ubiquitination & Proteasome degradation	-5.7
	MCODE_12	GO:0000209	protein polyubiquitination	-5.7
	MCODE_12	R-HSA-983169	Class I MHC mediated antigen processing & presentation	-5.4
	MCODE_13	R-HSA-8856825	Cargo recognition for clathrin-mediated endocytosis	-7.1
	MCODE_13	R-HSA-8856828	Clathrin-mediated endocytosis	-6.7
	MCODE_13	R-HSA-199991	Membrane Trafficking	-4.7
	MCODE_17	GO:0016584	nucleosome positioning	-9.5
	MCODE_17	GO:0031936	negative regulation of chromatin silencing	-9.3
	MCODE_17	GO:0031935	regulation of chromatin silencing	-8.4
	MCODE_18	M3008	NABA ECM GLYCOPROTEINS	-6.3
	MCODE_18	M5884	NABA CORE MATRISOME	-5.8
	MCODE_19	GO:0046580	negative regulation of Ras protein signal transduction	-8.1
	MCODE_19	GO:0051058	negative regulation of small GTPase mediated signal transduction	-7.9
	MCODE_19	R-HSA-6802949	Signaling by RAS mutants	-7.8

Table 3 List of the top 20 DEGs in T47D-RBMS3 vs. T47D cells from RNA-seq expression dataset.

Genes	Upregulated (FC)	Genes	Downregulated (FC)
RP11-9J18.1	3.626461	KRT16	-4.216279341
ANGPT1	3.269688	GREM2	-3.176657388
CLDN16	3.126702	HIST1H4D	-2.662753541
CRABP1	2.870269	CLEC3A	-2.479534141
GALNT5	2.759951	HPSE2	-2.350982455
CLDN1	2.484171	LA16c-312E8.5	-2.326493157
HBA1	2.387961	HIST1H1B	-2.135837845
LDHB	2.342741	TNFRSF11A	-2.0021068
FAM25A	2.103505	OR52E6	-1.908076923
EPHA6	2.062273	LMO4	-1.875299924

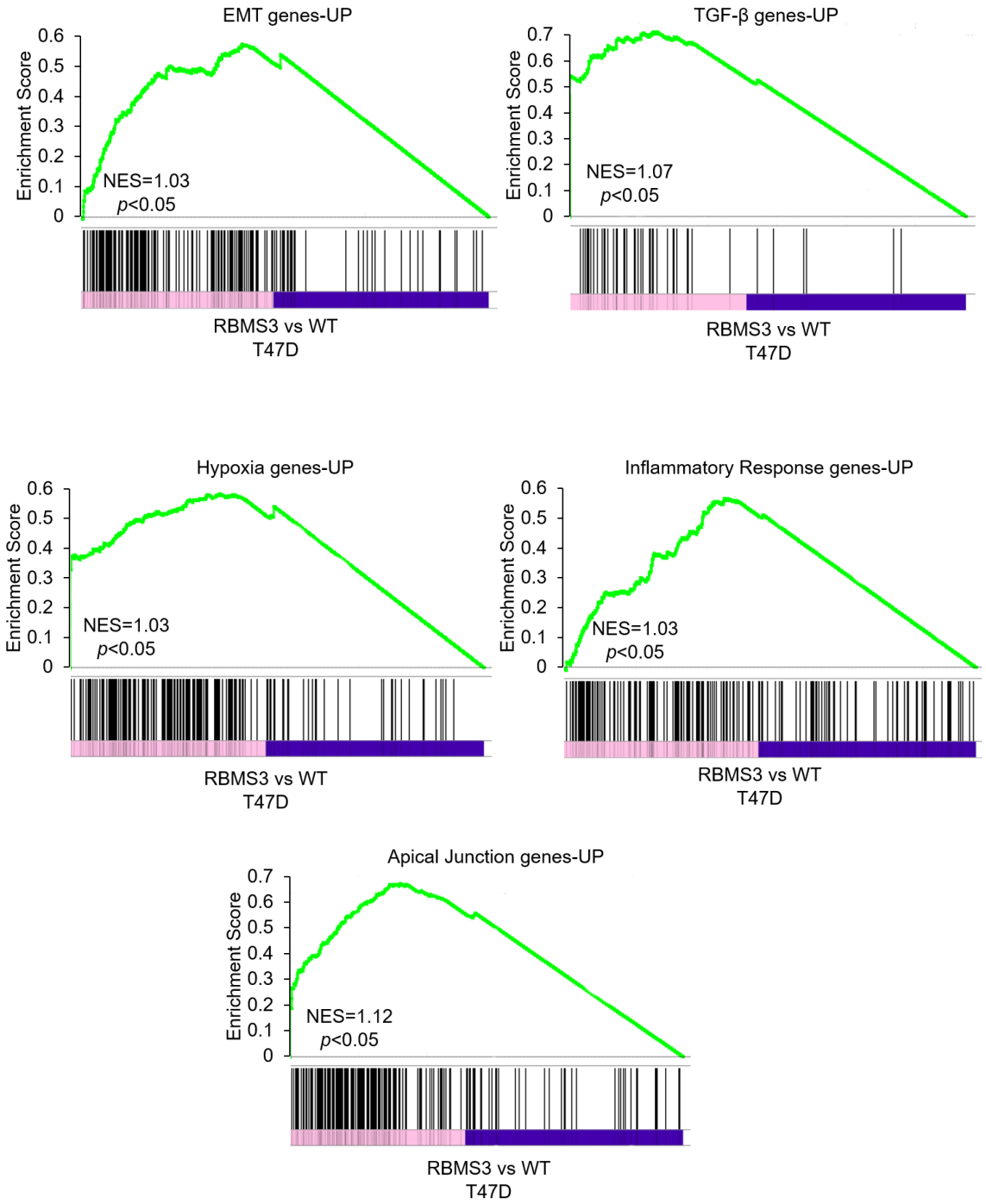


Figure 4.24 GSEA showing an enrichment of several pathways

Gene Set Enrichment Analysis (GSEA) showing an enrichment of the following pathways using DEG list of T47D-RBMS3 versus T47D control cells: EMT, inflammatory response, apical junction, hypoxia and TGFβ-related gene signatures. NES indicates normalized enrichment score.

Table 4 Selected pathways from GSEA analysis with highest ranked NES (>1.00) indicating high probability of positive correlation.

Gene Set Examples	NES	Correlation with Gene of Interest
Apical Junction	1.12	Positive
TGFβ Signaling	1.07	Positive
Unfolded Protein Response	1.05	Positive
PI3K-Akt-mTOR Signaling	1.04	Positive
Epithelial Mesenchymal Transition	1.03	Positive
Inflammatory Response	1.03	Positive
Hypoxia	1.03	Positive
Myc Targets v1	1.03	Positive
G2M Checkpoint	1.02	Positive
Wnt Signaling	1.02	Positive
Angiogenesis	1.01	Positive
Androgen Response	1.01	Positive
UV Response	1.01	Positive
Interferon-γ Response	1.01	Positive
TNF-α Signaling via NFκB	1.01	Positive
P53 Signaling	1.01	Positive
KRAS Signaling	1.01	Positive

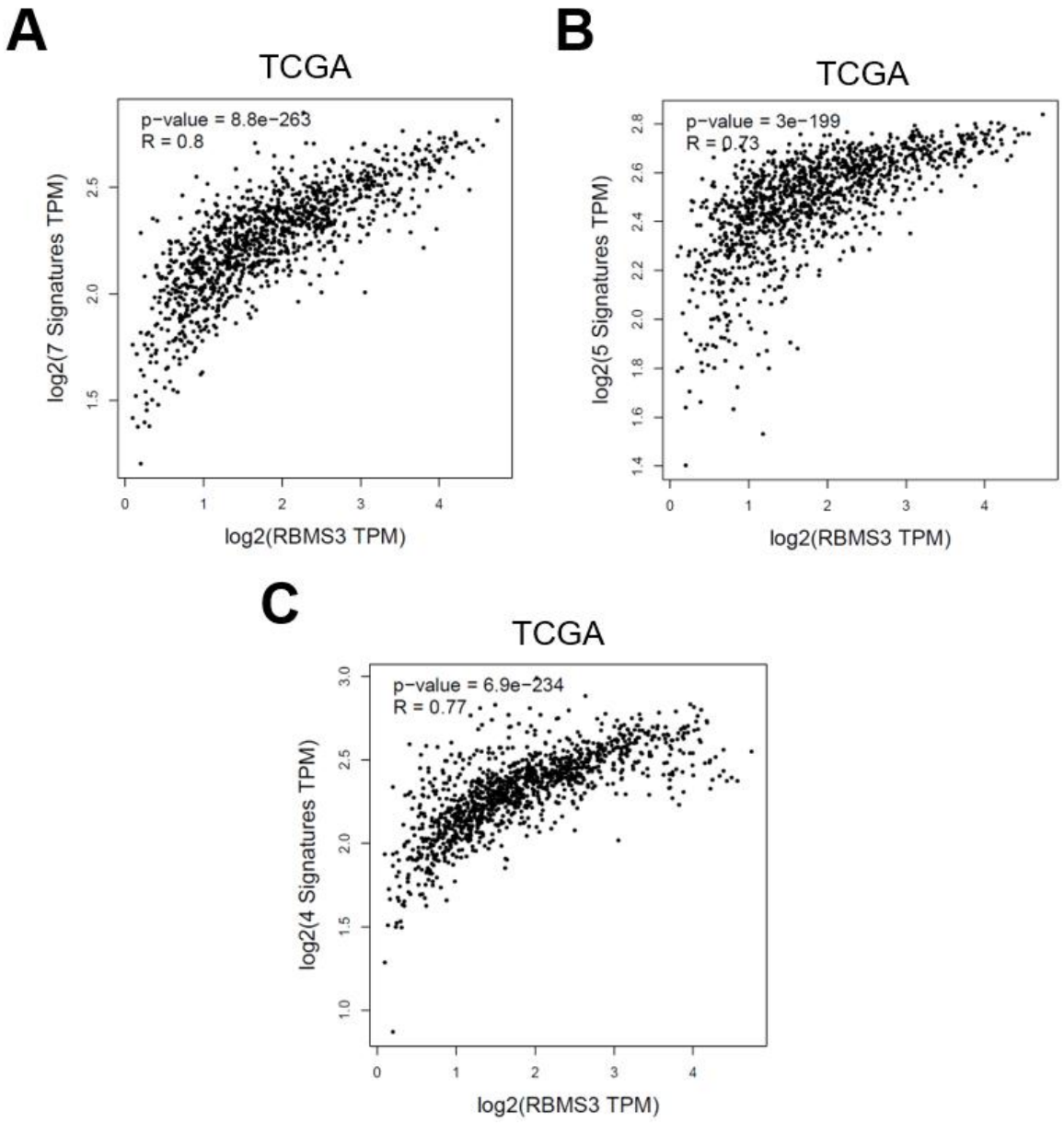


Figure 4.25 Correlation of RBMS3 and EMT/ TGFβ-related gene signatures
Correlation analysis was performed on the Gene Expression Profiling Interactive Analysis version 2 (GEPIA2) online analysis tool, which analyzes BRCA dataset from TCGA database. R denotes correlation efficient. TPM, transcript per million. a) Correlation between RBMS3 and seven EMT signatures (*SNAI1*, *SNAI2*, *TWIST1*, *TWIST2*, *ZEB1*, *ZEB2*, *VIM*). b) Correlation between RBMS3 and five TGFβ pathway component signatures (*TGFBR1*, *TGFBR2*, *SMAD2*, *SMAD3* and *SMAD4*). c) Correlation between RBMS3 and four TGFβ pathway targets (*PTGS2*, *CTGF*, *SMAD7* and *SNAI2*).

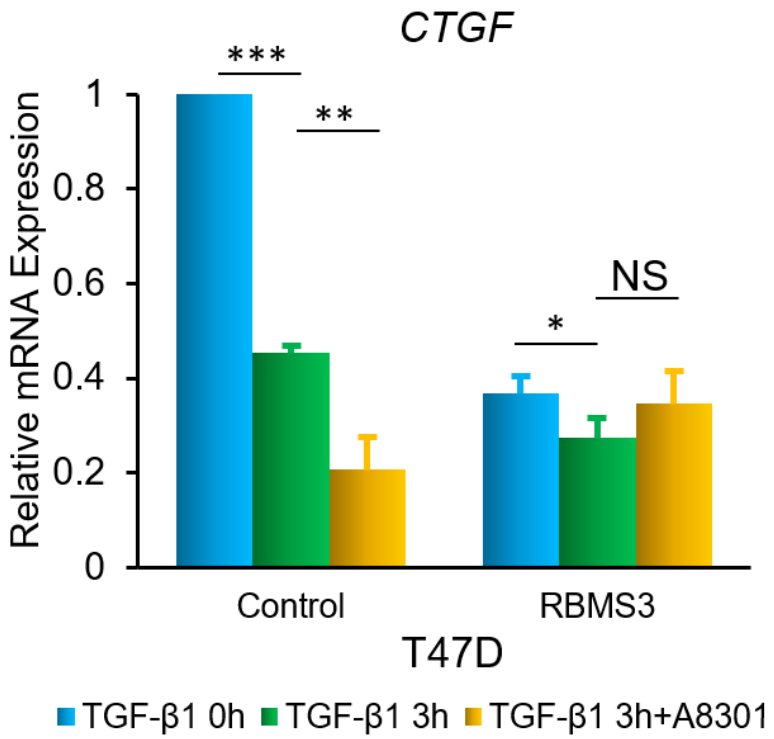


Figure 4.26 RBMS3 promotes downregulation of TGF-β1-regulated *CTGF* gene expression in T47D cells

Real-time PCR analysis of TGF-β downstream gene *CTGF* level in T47D cells stably expressed RBMS3 or control vector treated with TGF-β1 (10 ng/mL) or TGFβR1 inhibitor A8301 (5 μM). Data are shown as mean ± SD from two independent experiments in triplicates.

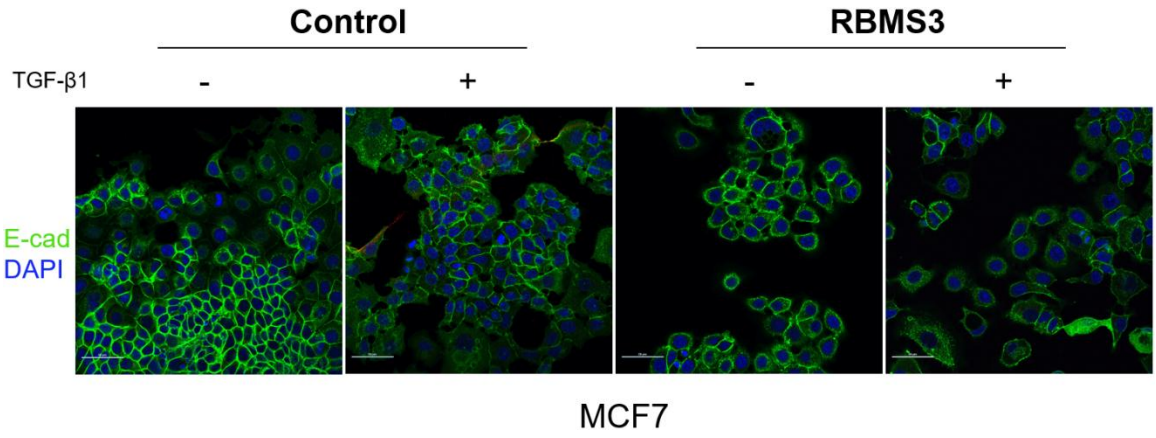


Figure 4.27 RBMS3 facilitates TGF-β1-mediated downregulation of E-cadherin

Representative immunofluorescence images of MCF7 cells stably expressing RBMS3 or control vector left untreated or treated with TGF-β1 (10 ng/mL) for 24 hours, and then immunostained with anti-E-cadherin antibody and DAPI. Scale bars, 40 μm.

4.6 RBMS3 alleviates TGF-β1-mediated cytostasis and apoptosis in luminal breast cancer cells

TGF-β signaling features a growth inhibitory effect at premalignant stages of breast cancer cells, but aggressive oncogenic activities at advanced malignant state [54, 150]. Long term treatment (~two weeks) of TGF-β1 induced massive cell death in MCF7 cells, but exerted significantly lower impact on MCF7-RBMS3 cells (Figure 4.28). Since RBMS3 posed little impact on MCF7 cell proliferation (Figure 4.6A), this data indicated the possibility of RBMS3 in suppressing TGF-β1-induced cell death. To confirm this data, MCF7 control and RBMS3-expressing cells were subjected to trypan blue exclusion assay. MCF7-RBMS3 cells showed significantly reduced dead cells compared to control cells starting at day 3 post treatment of TGF-β1 (Figure 4.29). To determine whether apoptosis is the cause of cell death induced by TGF-β1, cells were examined for biochemical and morphological markers of apoptosis using Propidium Iodide/Annexin V staining. After 24 hour TGF-β1 treatment, MCF7-RBMS3 and control cells showed similar baseline apoptotic population of 0.6% (data not shown). However, after 72 hours, TGF-β1 induced a drastic increase in the apoptotic cell population to 54.3% in control cells versus 36.7% in MCF7-RBMS3 cells (Figure 4.30). To study the effect of RBMS3 on the TGF-β1-mediated cytostatic effect on MCF7 cells, we examined cell cycle distribution after 24 hour TGF-β1 treatment. TGF-β1 induced an increase in the fraction of S phase cells and decrease in the fraction of G2/M phase MCF7 cells. Upon TGF-β1 stimulation, compared to MCF7 control cells, MCF7-RBMS3 showed significantly higher fraction of S phase cells and lower fraction of G0/G1 phase cells, suggesting that more cells entered into S phase and that RBMS3 seems to prevent the TGF-β1-mediated G1-S phase arrest (Figure 4.31). Upon 24 hours of TGF-β1 stimulation, MCF7 control cells showed upregulation of pro-apoptotic proteins Bad and cleaved Caspase 3, and downregulation of cell cycle protein CDC2 compared to untreated cells, suggesting increased apoptosis and decreased proliferation in the TGF-β1-treated MCF7 cells. Upon TGF-β1 stimulation, however, MCF7-RBMS3 cells

showed a reduction in pro-apoptotic protein Bax and unchanged levels of Bad, cleaved Caspase 3 and CDC2 compared to untreated cells, suggesting decreased apoptosis and unaffected proliferation in these cells (Figure 4.32). These data suggest that RBMS3 may alleviate TGF-β1-mediated apoptotic and cytostatic effects on MCF7 cells.

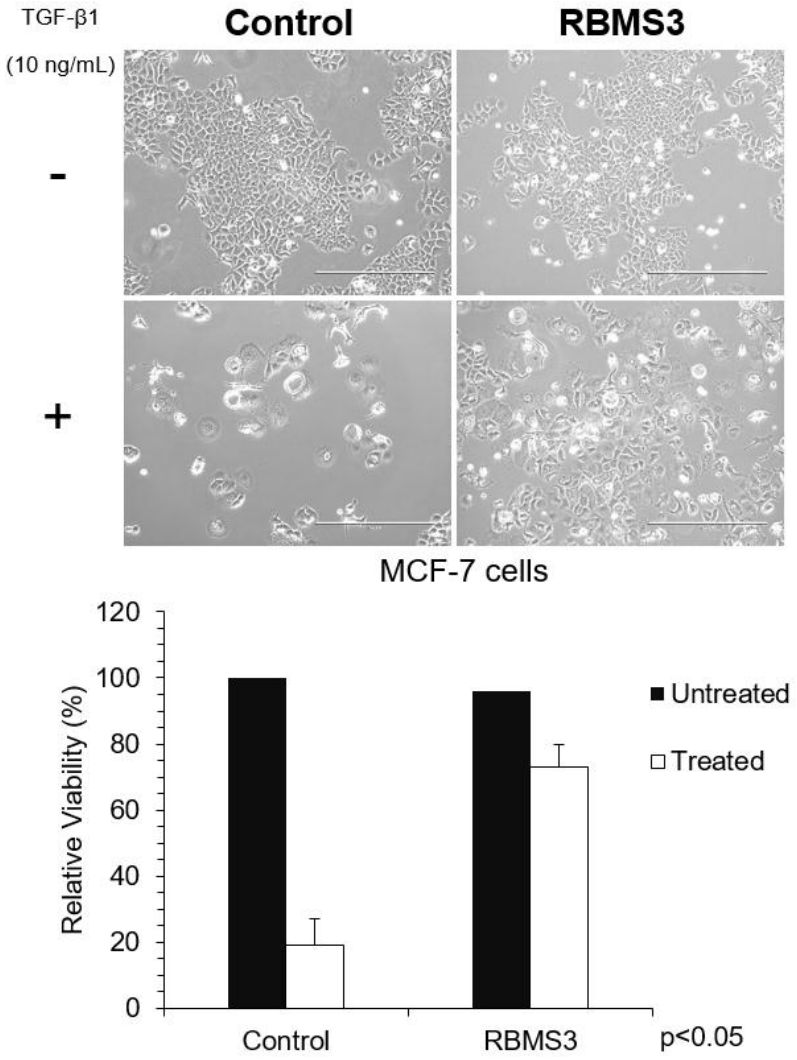


Figure 4.28 RBMS3 suppresses TGF-β1-induced apoptosis in MCF7 cells
Representative images of MCF7 cells stably expressed control or RBMS3 vector followed by administration of EGF (10 ng/mL) and TGF-β1 (10 ng/mL) for two weeks. Relative viability is measured by cell counting and data are plotted using untreated MCF7 cells as control. Scale bars, 400 μm. Data are shown as mean ± SD from three independent experiments. *** $p < 0.001$ when TGF-β1-treated MCF7-RBMS3 group is compared with TGF-β1-treated MCF7 group.

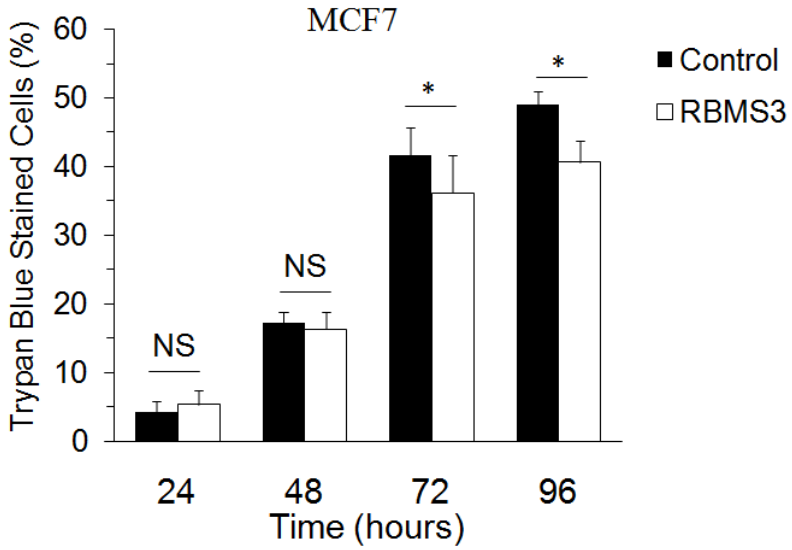


Figure 4.29 RBMS3 prevents TGF-β1-induced cell death in MCF7 cells

Cells were treated with TGF-β (10 ng/mL) for indicated period of time, typsinized and counted for nonviable cells with trypan blue. Data were represented as mean ± SD from three independent experiments. (NS indicates nonsignificant, * indicates $p < 0.05$)

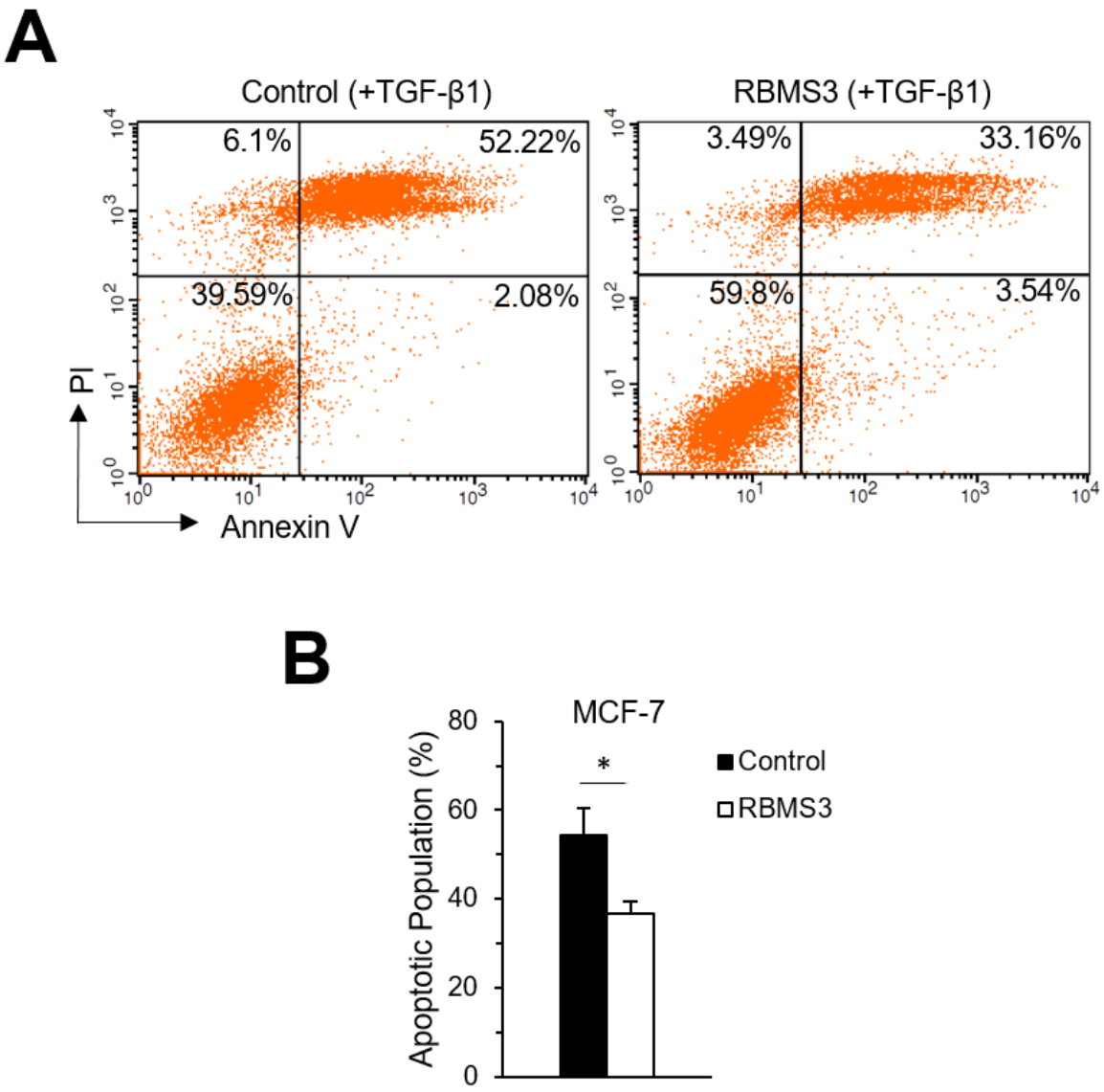


Figure 4.30 Annexin V/PI staining of MCF7-RBMS3 cells treated with TGF-β1
(A) Representative FACS images from MCF-7 cells expressed control or RBMS3 vector with TGF-β1 treatment for 72 hours using Annexin V-FITC Early Apoptosis Detection Kit. Annexin V-FITC conjugated protein binds to cell surfaces expressing phosphatidylserine, an early apoptosis marker. Cells stained with propidium iodide (PI), a non-cell-permeable DNA dye, indicate necrotic cells. Cells stained with both Annexin V-FITC and PI demonstrate late stage apoptosis and early necrosis.

(B) Quantification of apoptotic population between MCF-7 cells expressed control or RBMS3 vector under TGF-β1 treatment is shown. Data are shown as mean ± SD from two independent experiments. ***p value < 0.005

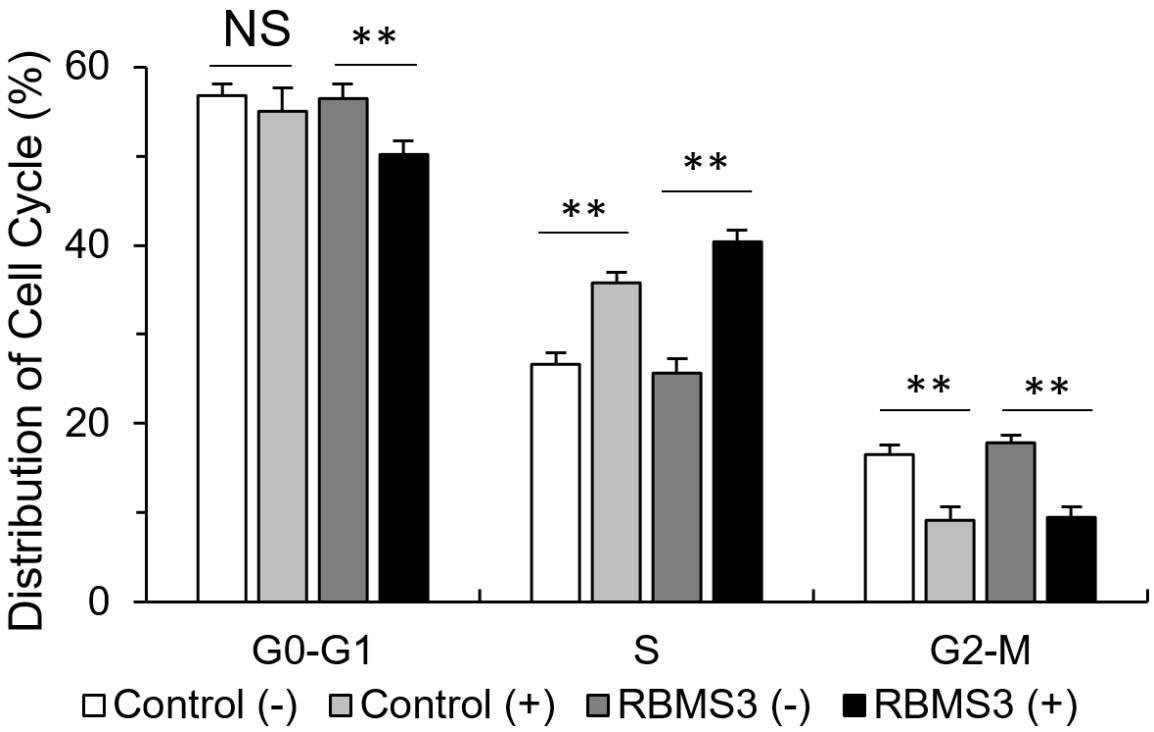


Figure 4.31 RBMS3 facilitates TGF-β1-mediated increase of S-phase population
Statistical analysis of Cell cycle distribution (G0/G1 phase, S phase or G2/M phase) of MCF7 cells stably expressed RBMS3 or control vector. Cells were starved overnight and then left untreated or treated with TGF-β1 for 24 hours.

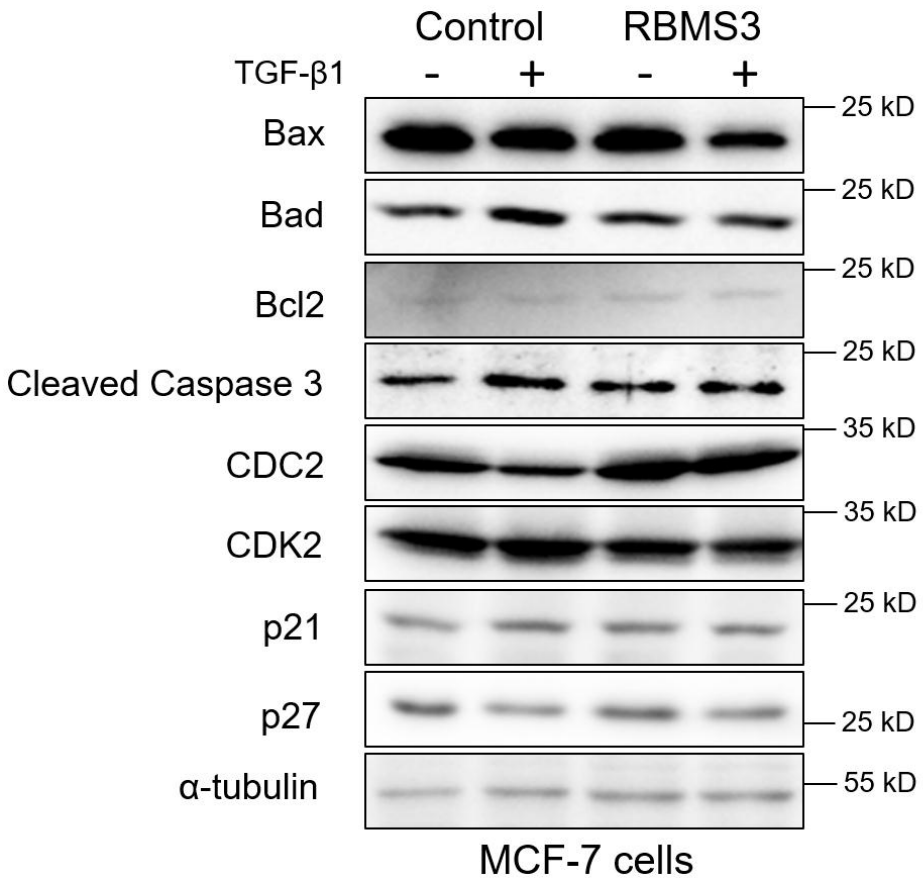


Figure 4.32 RBMS3 regulates TGF-β1-mediated apoptosis and cell cycle markers
Western blot analysis of apoptosis regulators (Bax, Bad, cleaved Caspase 3 and Bcl2) and cell cycle regulators (CDC2, CDK2, p21 and p27) in MCF7 cells stably expressed control or RBMS3 vector in the presence or absence of TGF-β1 (10 ng/mL) for 24 hours. α-tubulin was loaded as control.

4.7 Loss of RBMS3 Disrupts Smad-dependent TGF-β Signaling

On the other front, RBMS3 depletion significantly downregulated TGF-β effectors *Smad2* and *Smad3* (Figure 4.33). Furthermore, loss of RBMS3 suppressed TGF-β-regulated genes (Figure 4.34). Of note, *CTGF* and *PTGS2* are recognized to be drivers of breast cancer bone metastasis [151]. RBMS3 depletion led to significant decreases in the total level of Smad2, Smad3 and Smad4 protein and the phosphorylation level of Smad2 and Smad3 in BT549 cells. We also noticed that EMT transcription factors Snail, Slug and Twist are significantly reduced upon loss of RBMS3 (Figure 4.35). These data further confirm that RBMS3 is involved in TGF-β-mediated EMT.

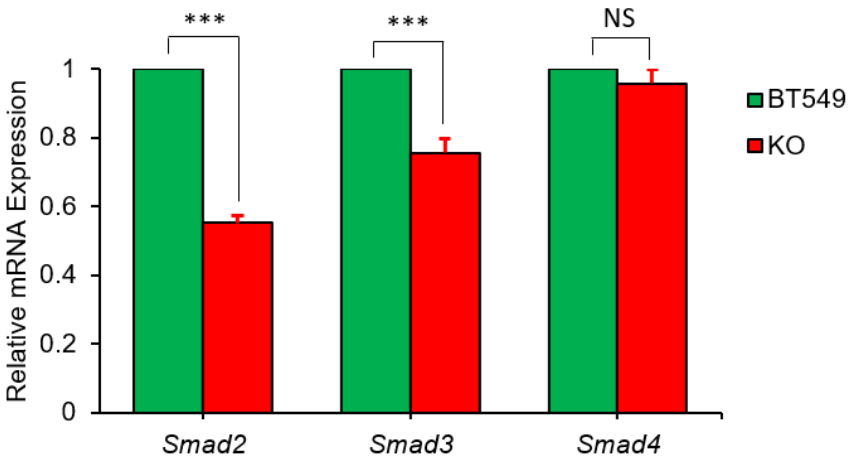


Figure 4.33 Loss of RBMS3 downregulates *Smad2* and *Smad3*
Real-time PCR analysis of the mRNA expression of TGF-β signaling mediators *Smad2*, *Smad3* and *Smad4* in BT549 cells with or without loss of RBMS3. ****p* <0.001 and NS stands for statistically non-significant. Data are shown as mean ± SD from three independent experiments in triplicates.

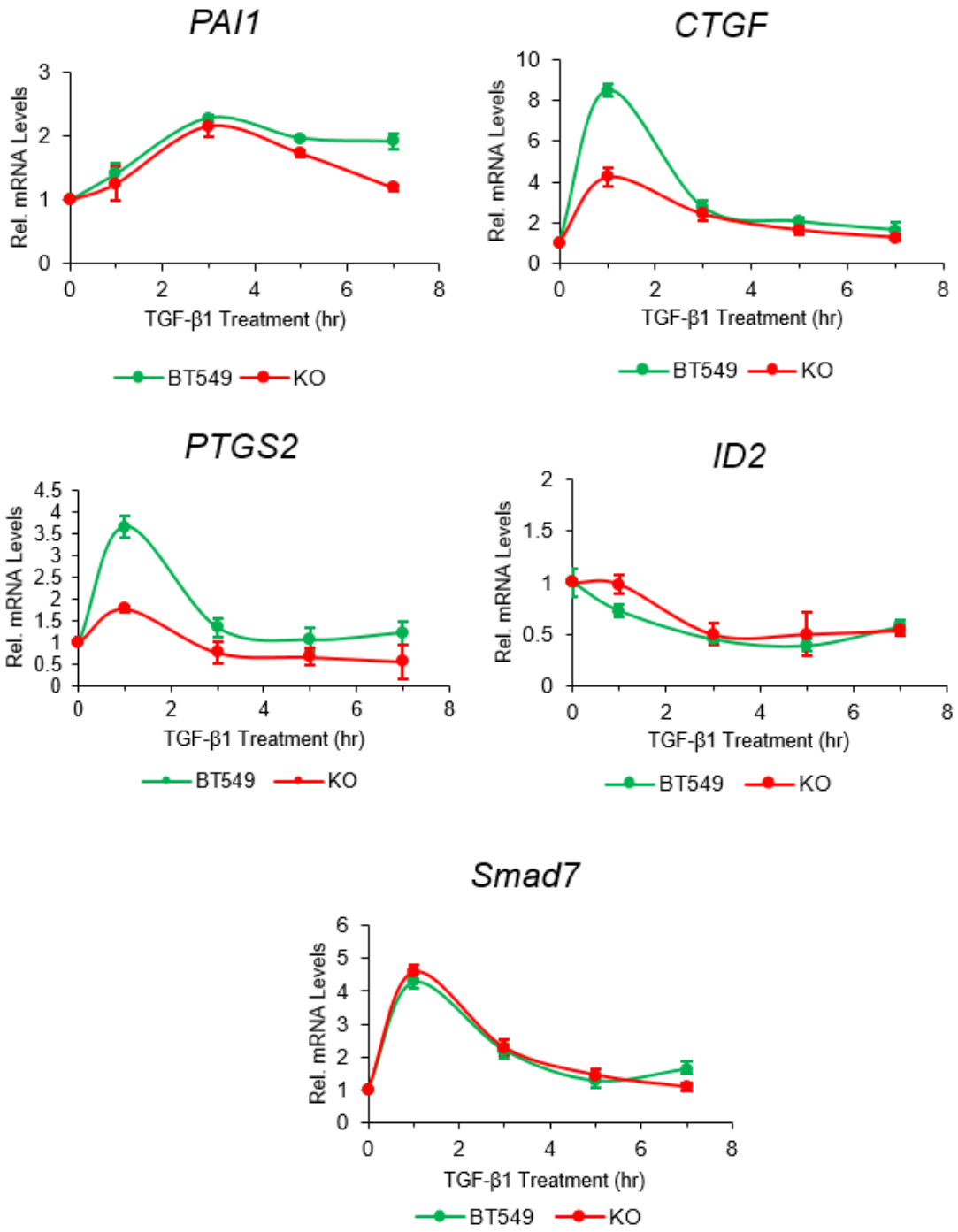


Figure 4.34 Loss of RBMS3 downregulates TGFβ1-regulated genes

Real-time PCR analysis of TGF-β downstream gene *PAI1*, *CTGF*, *ID2*, *PTGS2* and *Smad7* levels in BT549 cells with or without loss of RBMS3 treated with TGF-β1 for 0, 2, 4, 8 hours. Data are shown as mean ± SD from three independent experiments in triplicates.

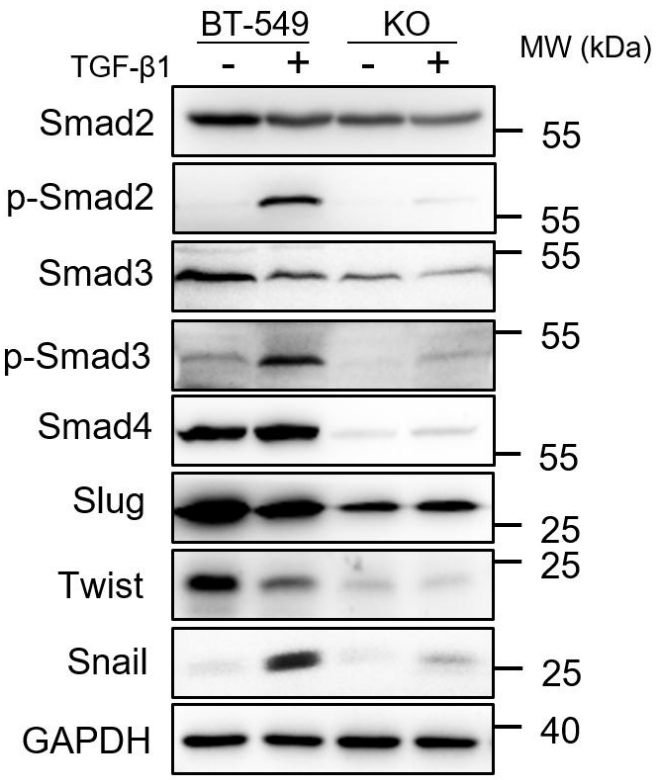


Figure 4.35 Loss of RBMS3 reduces TGF-β1 effector Smad proteins and EMT transcription factors
Western blot analysis of TGF-β signaling mediators (Smad2, p-Smad2, Smad3, p-Smad3 and Smad4) and mesenchymal markers (Slug, Twist, Snail) in BT549 cells with or without loss of RBMS3 left untreated or treated with TGF-β1 at 10 ng/mL. GAPDH was loaded for normalization.

4.8 RBMS3 regulates TGF-β signaling by stabilizing Smad transcripts

Once activated by TGF-β ligands, TGF-β signaling is mediated through Smad and non-Smad pathways to regulate transcription, translation, post-translational modifications, protein synthesis and RNA biogenesis [56]. In the canonical TGF-β pathway, TGF-β ligand binds to the type 2 TGF-β receptor (TGFBR2), which recruits the TGFBR1. The receptors dimerize and undergo autophosphorylation, allowing for the recruitment and phosphorylation of Smad2/3 by TGFBR1. The activated Smad2/3 dissociate from the plasma membrane, hetero-oligomerize with Smad4 and translocate into the nucleus to mediate gene expression/repression. Since protein levels changes of Smad2/3 were

observed in cells transfected with CRISPR/Cas9-RBMS3 or ectopic RBMS3 in comparison with mock-treated controls, we asked whether RBMS3 regulates Smad2/3 transcripts. To test this possibility, we first examined the subcellular localization of RBMS3. As expected, we found that RBMS3 was mainly localized in the cytoplasm, suggesting roles in RNA metabolism (Figure 4.36). Next, we asked whether RBMS3 physically interact with Smad2/3 transcripts. Lentiviral vectors of strep-tagged full-length RBMS3 (RBMS3-Strep) and strep-tagged mutant RBMS3 lacking the two putative RNA-binding domains (dRRM-RBMS3-Strep) were generated, respectively (Figure 4.37A). RNA immunoprecipitation assay followed by RT-PCR was performed using cell extracts from MDA-MB-231 cells that were induced to express RBMS3-Strep or dRRM-RBMS3-Strep (Figure 4.37B). We found significant enrichment of Smad2/3 transcripts in MDA-MB-231 cells expressing RBMS3-Strep compared to dRRM-RBMS3-Strep, indicating physical interaction between RBMS3 and Smad2/3 transcripts (Figure 4.38). RNA-binding proteins are known to regulate gene expression at post-transcriptional levels, including mRNA stability. In zebrafish prechondrogenic crest cells, RBMS3 may regulate the stability of *Smad2* transcript to control the pool of protein available for signaling [136]. To explore whether this mechanism applies to TNBC cells, the mRNA stability of *Smad* transcripts was measured. Upon treatment with transcription inhibitor actinomycin D (5 μg/ml), knockout of RBMS3 significantly decreased the half-lives of *Smad2/3/4* transcripts in BT-549 cells (Figure 4.39). Altogether, our results suggest that RBMS3 post-transcriptionally modulates TGF-β signaling by stabilizing *Smad2/3/4* mRNAs.

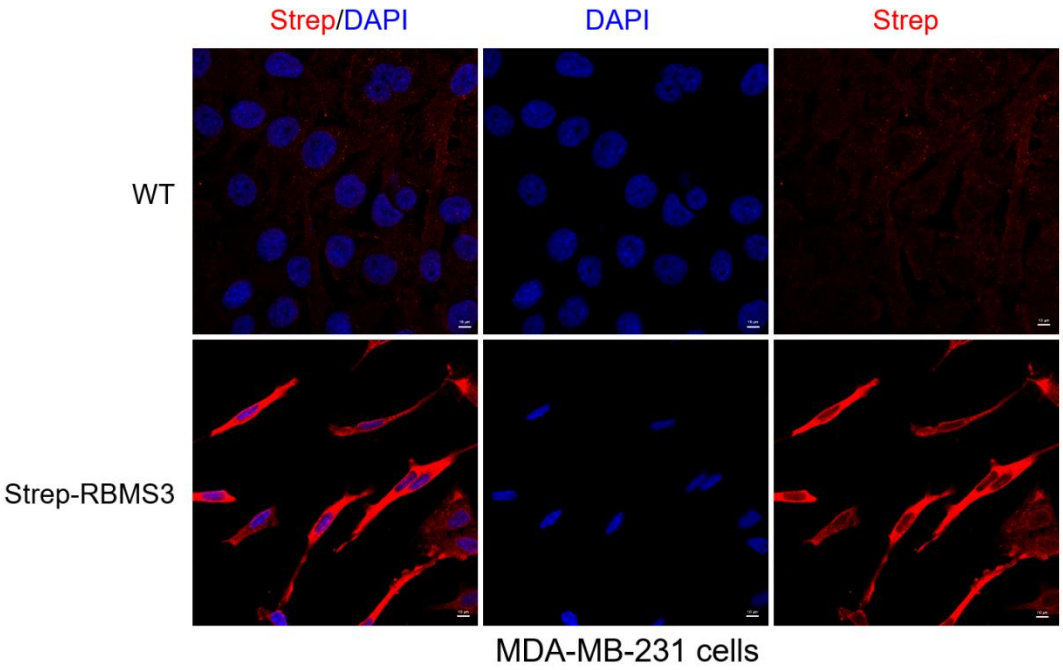


Figure 4.36 RBMS3 is mainly localized in the cytoplasm

Representative immunofluorescence images of MDA-MB-231 cells stably expressed control or Strep-RBMS3 vector immunostained by anti-Strep antibody, visualized by goat anti-mouse conjugated with Alexa fluor 568 under confocal microscope.

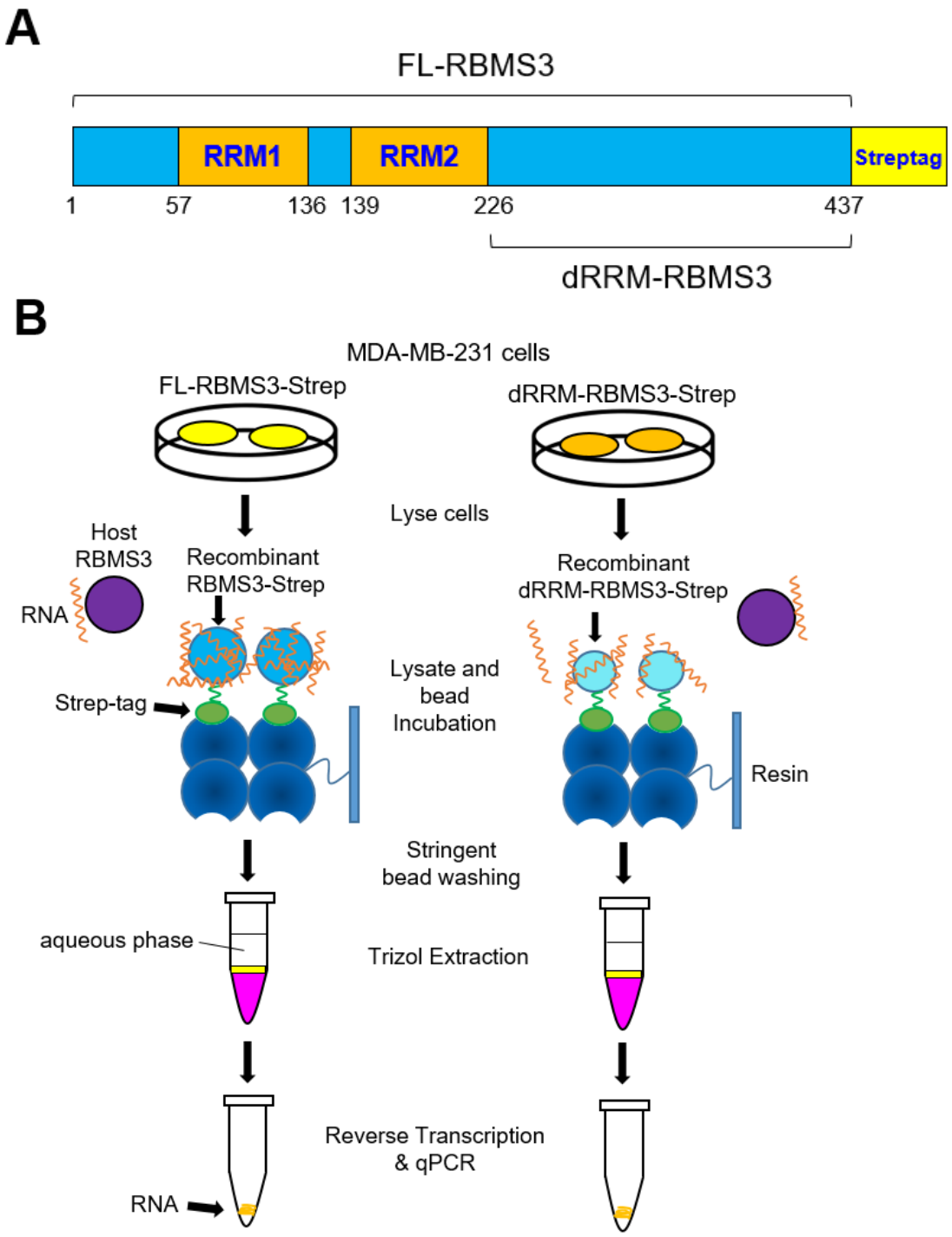


Figure 4.37 Schematic of constructs and RNA-Immunoprecipitation Procedure
(A)Graphic representation of FL-RBMS3-Strep (full length) and dRRM-RBMS3-Strep (RRM motif-deficient) constructs. RRM denotes RNA recognition motif (RRM). Numbers of amino acids denote the start and end of each protein segment or motif.

(B) Graphic representation of general RNA-Immunoprecipitation workflow. Detailed protocol is described in Materials and Methods Chapter below.

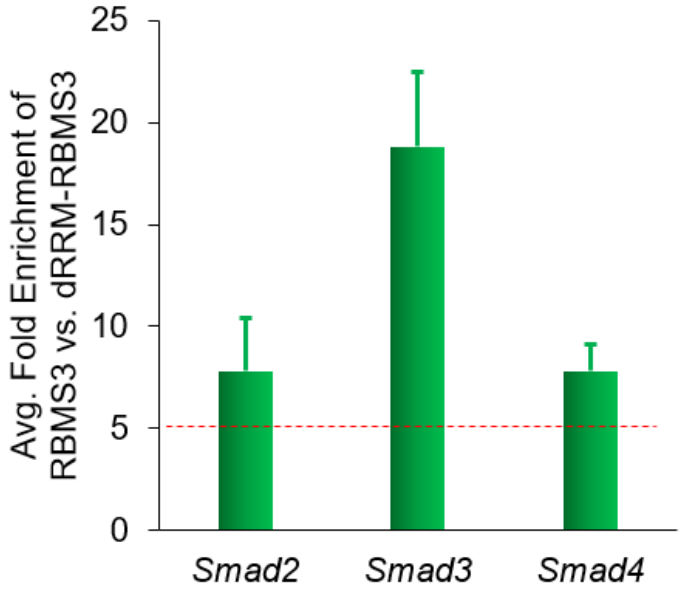


Figure 4.38 RNA-Immunoprecipitation assay showing enrichment of *Smad2/3/4* transcripts

Real-time PCR analysis of *Smad* transcripts enriched (fivefold as cutoff, red dashed line) for RBMS3-Strep (n=3 experiments). Data are shown as mean ± SD from three independent experiments. Other transcripts that are insignificantly enriched are not shown here.

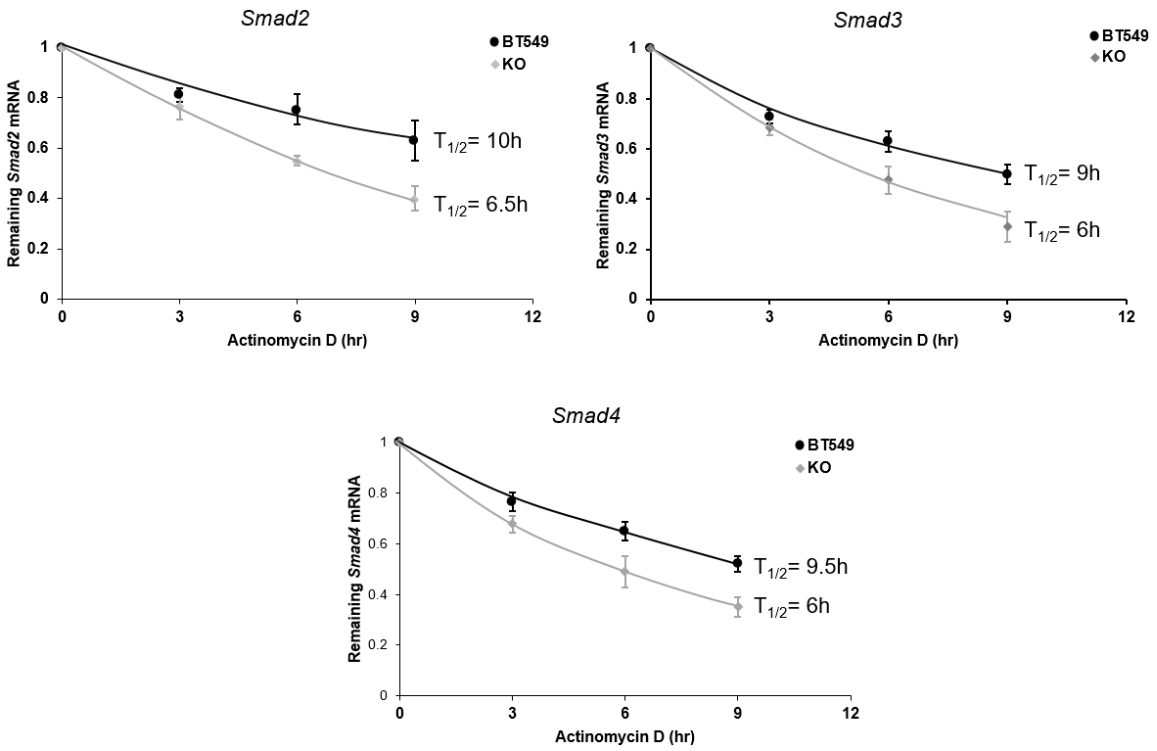


Figure 4.39 RBMS3 regulates *Smad2/3/4* mRNA stability

Real-time PCR analysis of remaining *Smad* transcripts (normalized to the level of actin transcript) in BT549 cells with or without RBMS3, and plotted along with time after Actinomycin D (5 μg/mL) treatment. T1/2 represents relative half-life of the specific *Smad* mRNA in the presence or absence of RBMS3.

4.9 RBMS3 is negatively correlated with ESRPs expression

Luminal BC cells are well-known to be rather difficult to become metastatic, largely due to presence of intrinsic master regulators that govern the identity of the cell type. ESRP1 and ESRP2 (ESRP1/2) are splicing regulators that govern epithelial isoforms of proteins during mammalian development. ESRP1/2 are essential for a range of epithelial cell properties, such as cytoskeletal dynamics, cell motility, cell-cell junctions and pathways involved in EMT [152]. Importantly, loss of ESRP1 and ESRP2 disrupts the ability to form a proper epithelial layer, concomitant with increased motility and expression of invasive markers, suggestive of acquisition of mesenchymal phenotypes. Therefore,

ESRPs are known to suppress EMT in various cancers. Our microarray dataset suggested *Esrp1* and *Esrp2* are overexpressed in luminal BC cells but repressed in BLBC cells, which is consistent with western blotting results (Figure 4.40-42). While ESRP1 contains 4 RRMs, a DnaQ-like 3'-5' exonuclease domain superfamily (DEDD) and a Domain of Unknown Function (DUF), ESRP2 contains 3 RRMs and a DEDD (Figure 4.43). To better understand the roles of ESRPs and RBMS3, we generated stable ESRP knockout clones and RBMS3 overexpression clones and subjected them to functional assays and RNA-seq. Western blotting suggested successful establishment of these clones (Figure 4.44-45). By appearance, T47D-ESRP1 KO cells are more stretched and edged, whereas T47D-ESRP2 KO cells appear more rounded and organized than control cells. T47D-ESRP DKO cells attach loosely to dishes, instead, they tend to aggregate to form clusters or grow on stacks. Surprisingly, T47D-ESRP DKO-RBMS3 cells regains the ability to attach to dishes, an indication that RBMS3 is involved in cell adhesion and spreading. None of the cell lines showed a typical mesenchymal appearance of a spindle-like shape, even in the presence of TGF-β (Figure 4.46). Despite of this, we asked whether RBMS3 contribute to the EMT process by ectopically expressing an EMT transcription factor Slug in each of the T47D-derived clones. Indeed, at day 9 post transfection, we noticed that Slug overexpression in T47D-ESRP DKO-RBMS3 cells resulted in ~ 12-fold increase in spindle-like shaped cells than Cas-9 expressing control cells, compared to the 3 ~ 4 fold increase in cells with loss of ESRP1/2 or RBMS3 expression (Figure 4.47). This suggests that RBMS3 may predispose the cells to a state that is prone to EMT. Moreover, long term expression (over 10 days) of EMT transcription factors Slug or Snail led to significant apoptosis in T47D cells with or without loss of ESRP1 or ESRP2. Overexpression of Twist in T47D cells with or without loss of ESRP1 also induced apoptosis, but not in cells with loss of ESRP2. Interestingly, loss of both ESRPs almost completely abolished apoptosis mediated by Slug, Snail or Twist (data not shown). The effects of TGF-β, Snail, Slug and Twist on the apoptosis and appearance of these cells lines are summarized in Table 5.

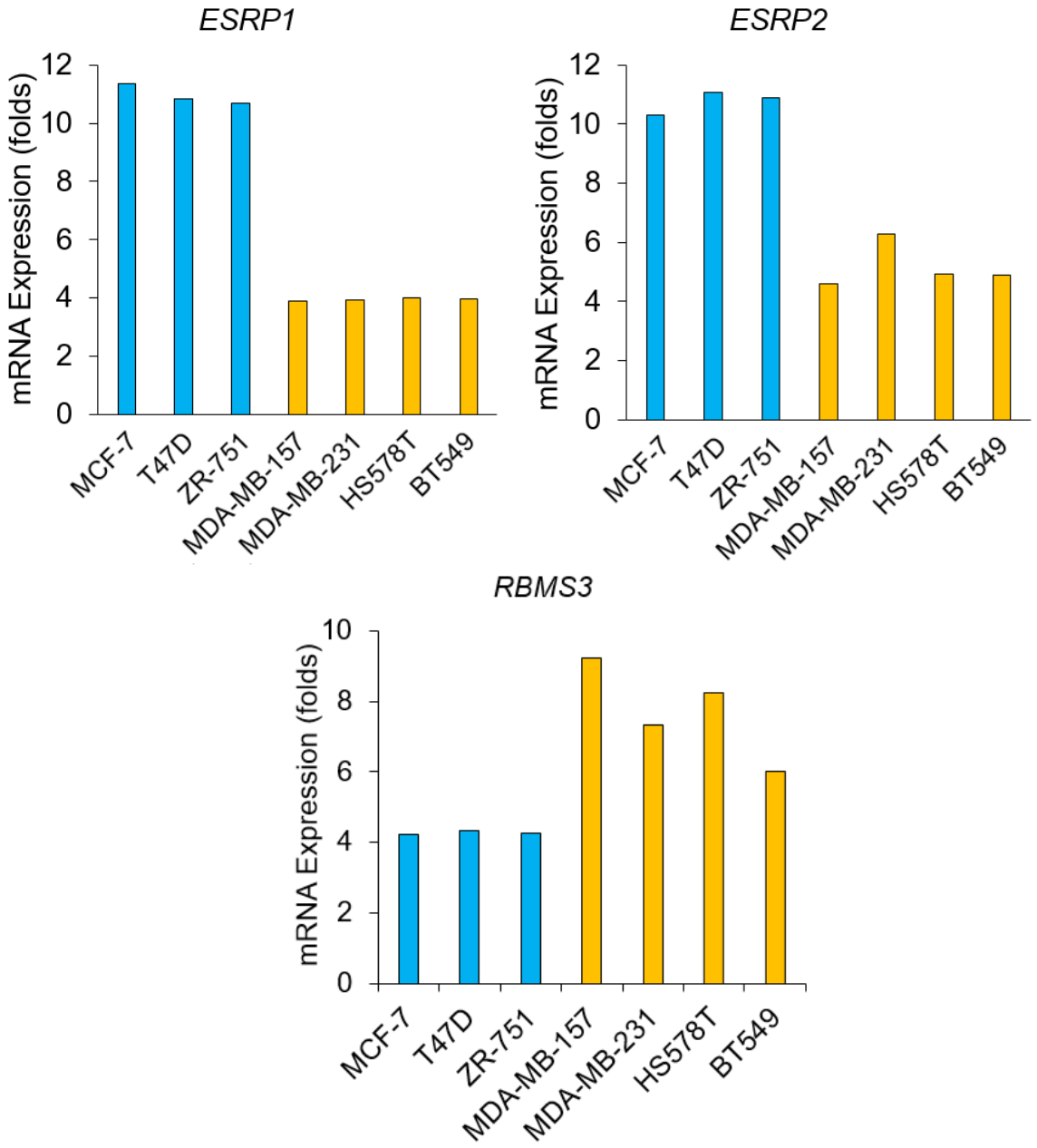


Figure 4.40 *Rbms3* is inversely correlated with *ESRP1* and *ESRP2* in TNBC cells compared to luminal BC cells (RNA-seq)

Data from CCLE RNA-seq expression dataset contained 3 luminal (MCF7, T47D and ZR-751) and 4 basal-like/triple negative (MDA-MB-157, MDA-MB-231, Hs578T and BT549) BC cell lines.

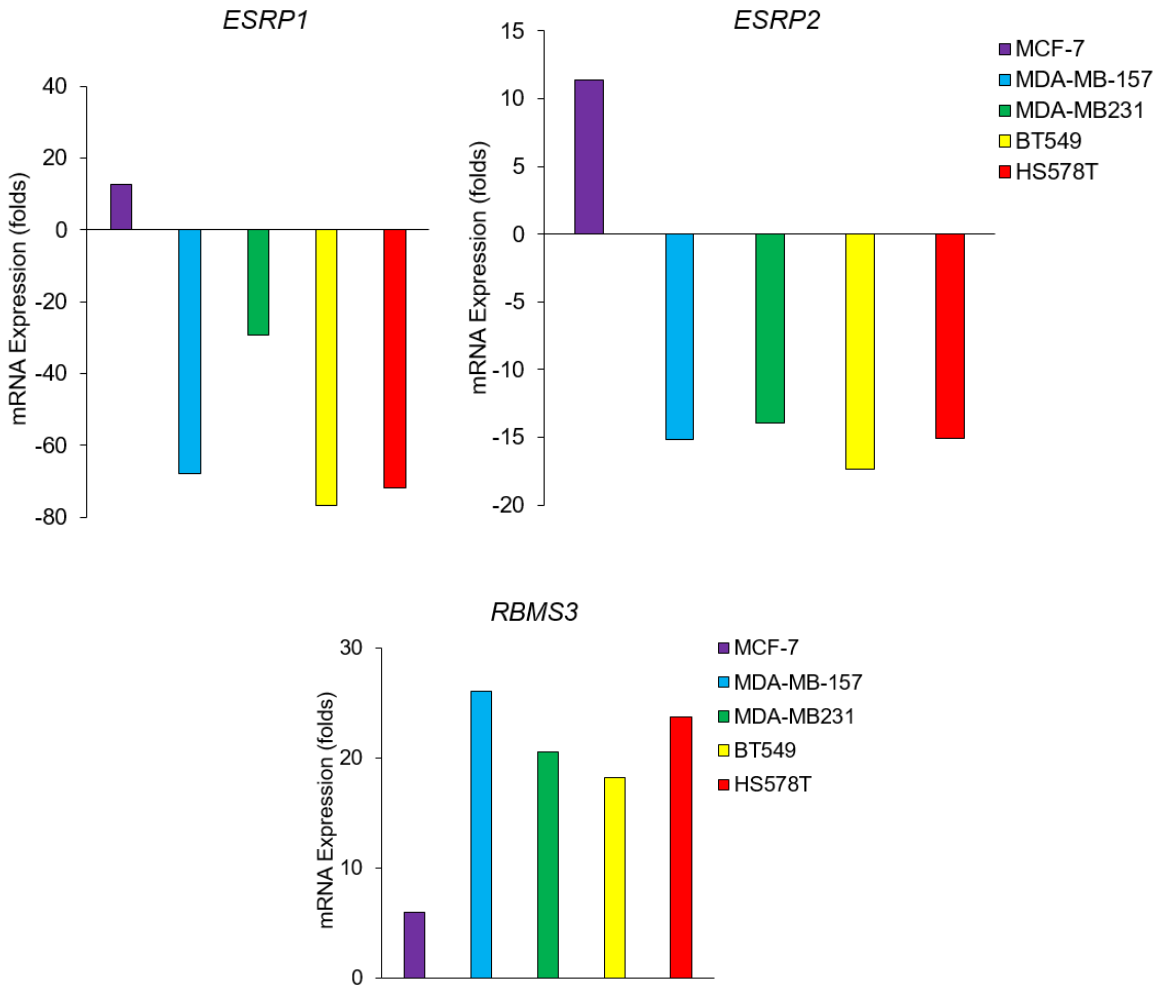


Figure 4.41 *Rbms3* is inversely correlated with *ESRP1* and *ESRP2* in TNBC cells compared to luminal BC cells (microarray)

Microarray Data dataset from our group contained luminal MCF7 cells and 3 basal-like/triple negative (MDA-MB-157, MDA-MB-231, Hs578T and BT549) BC cell lines.

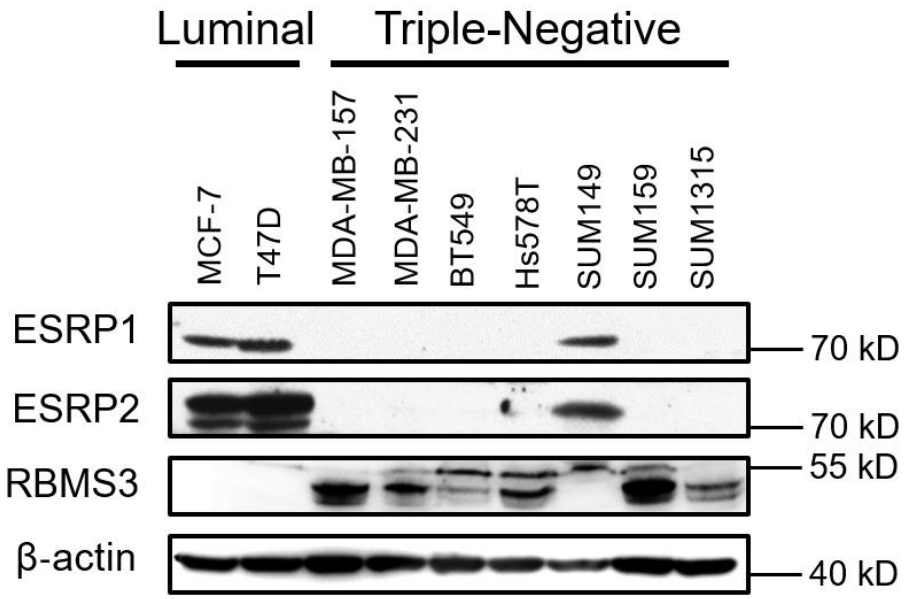


Figure 4.42 RBMS3 is inversely correlated with ESRP1 and ESRP*2* expression in TNBC cells compared to luminal BC cells

Western blot analysis contained 2 luminal cell lines (MCF7 and T47D) and 7 TNBC cell lines (including inflammatory BC cells SUM149 and 6 BLBC cells MDA-MB-157, MDA-MB-231, Hs578T, BT549, SUM159 and SUM1315).

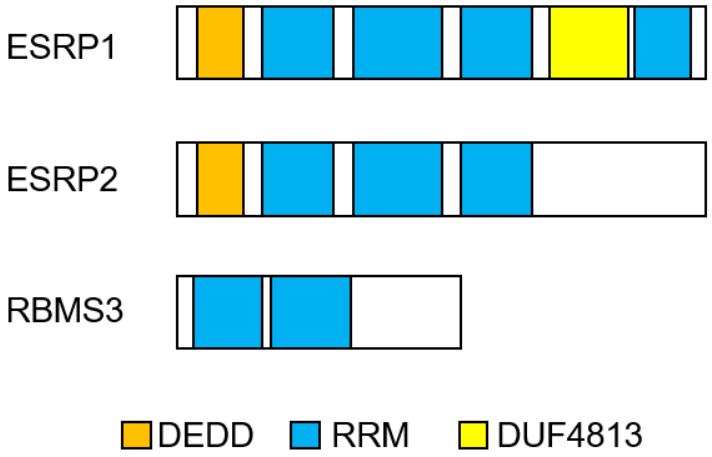


Figure 4.43 Schematic of the domain organization of ESRP1, ESRP2 and RBMS3

DEDD: DnaQ-like 3'-5' Exonuclease Domain superfamily; RRM: RNA Recognition Motif; DUF: Domain of Unknown Function. RBMS3 contains two RRMs that are known to bind single-stranded RNAs.

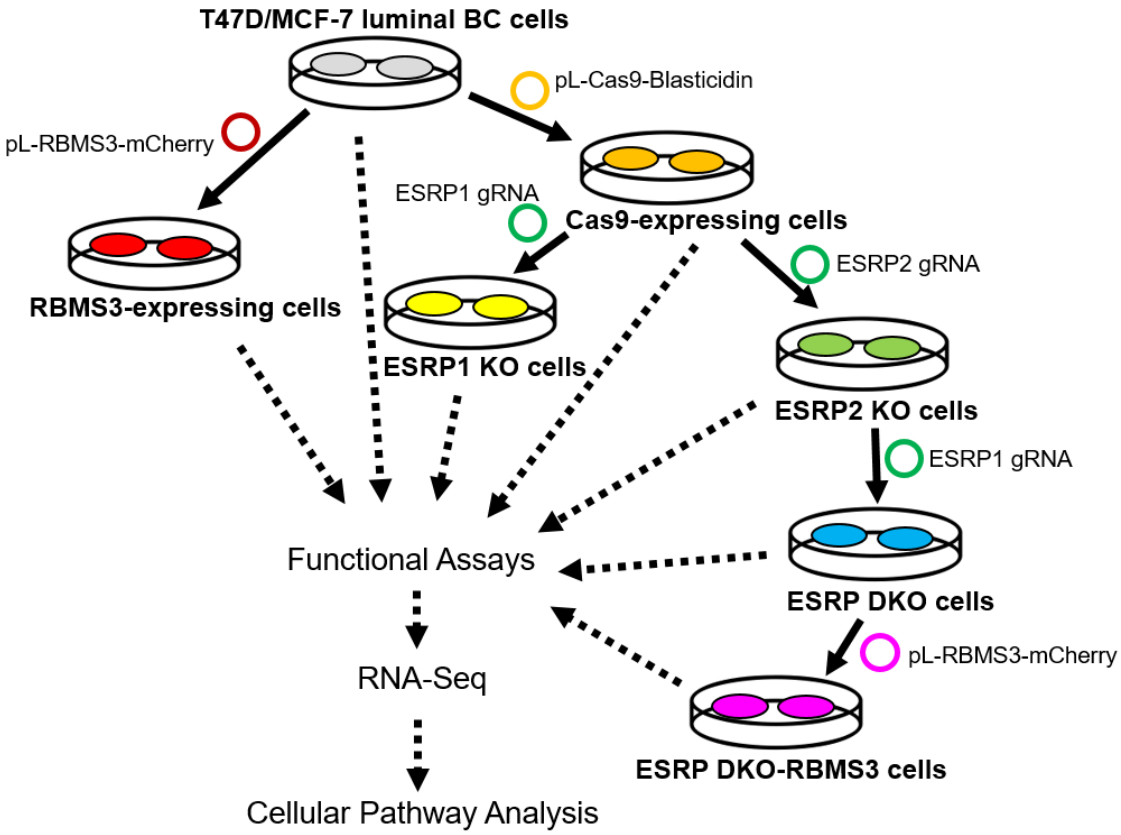


Figure 4.44 Schematic workflow for establishment of RBMS3 overexpression and ESRP knockout cell lines

All knockout cell lines were established on the Cas9-expressing cells. Each transfected cell line was either selected by corresponding drug for at least two weeks or cell sorting. All knockout cells were screened for single clones. A total of 14 cell lines were subjected to functional assays and RNA-seq.

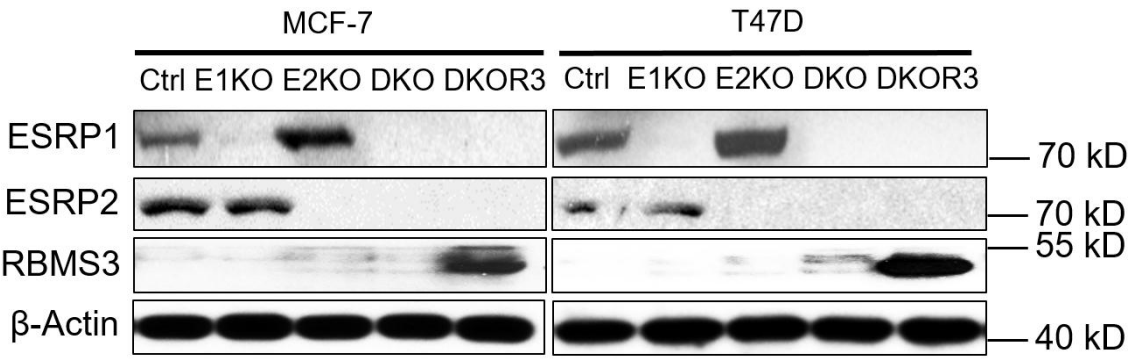


Figure 4.45 Establishment of ESRP-KO and RBMS3-OE cell lines

Western blot analysis of RBMS3 overexpression and ESRP knockout cell lines. (Ctrl denotes control/parental cells; E1KO denotes ESRP1 knockout; E2KO denotes ESRP2 knockout; DKO denotes ESRP1 and 2 double knockout; DKOR3 denotes ESRP 1 and 2 double knockout combined with RBMS3 overexpression)

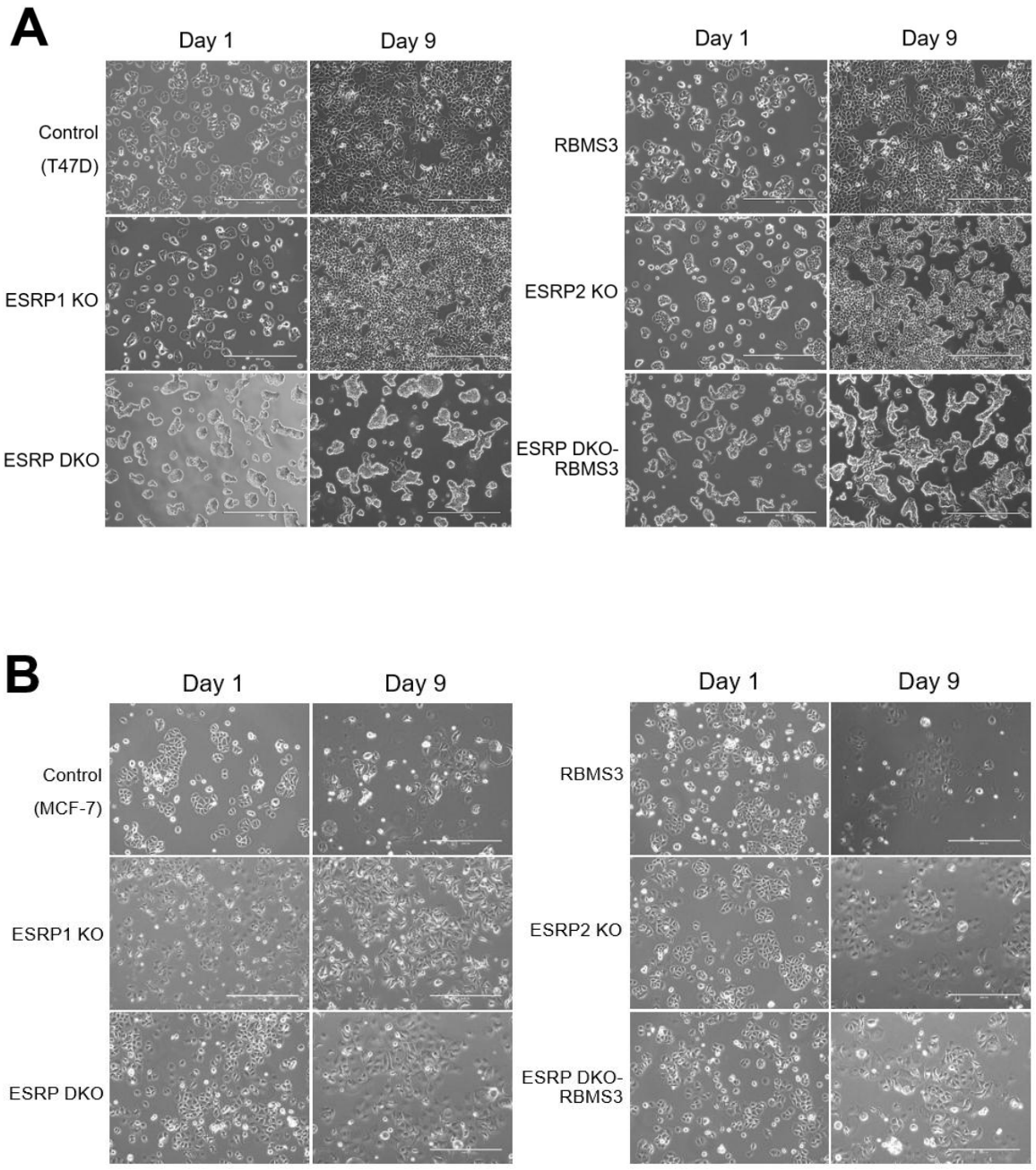


Figure 4.46 The effect of TGF- β1 on MCF7-derived or T47D-derived cells

Combination of RBMS3 expression and TGF- β1 treatment (10 ng/mL) does not contribute to typical mesenchymal appearance in MCF7 or T47D cells with loss of ESRP1 and ESRP2

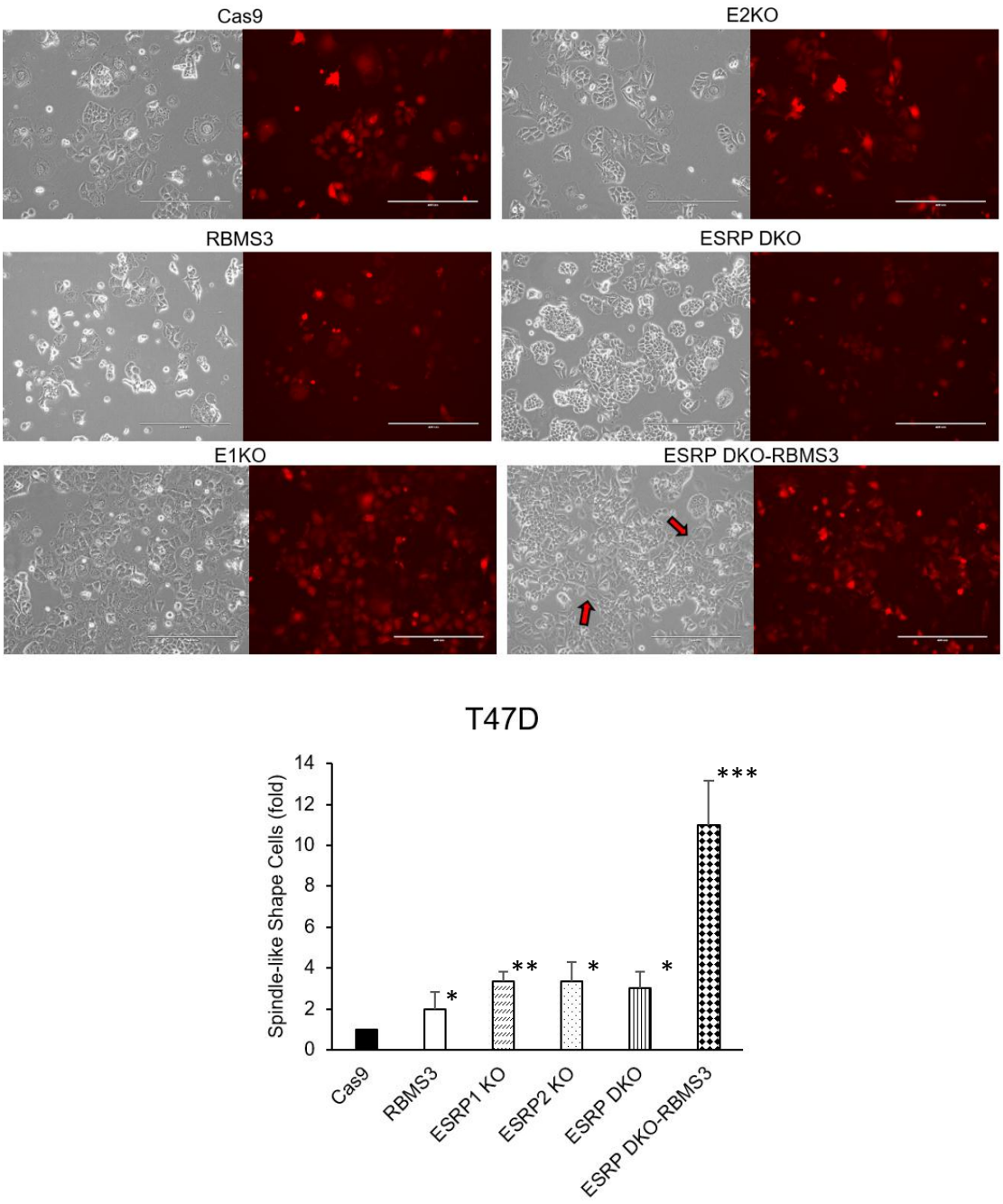


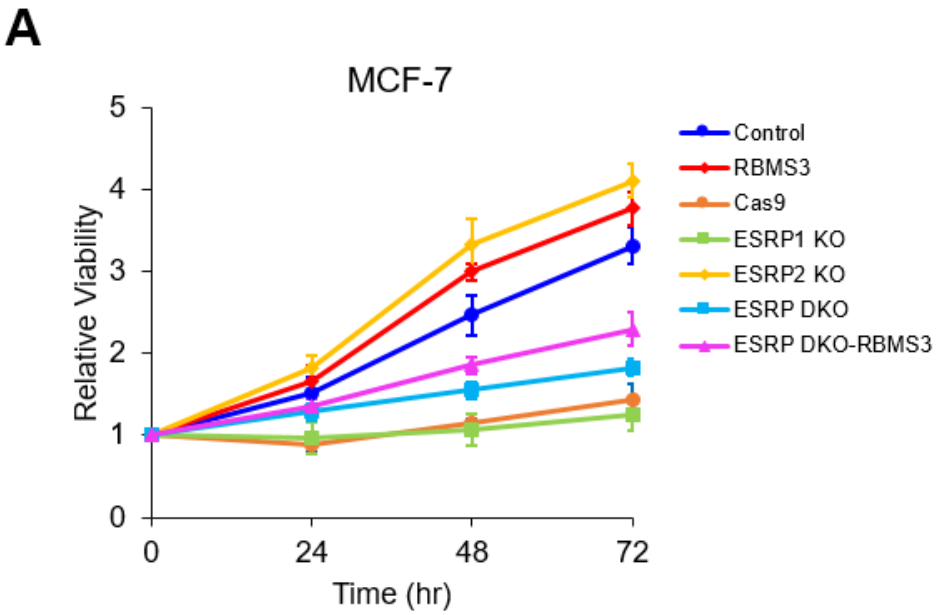
Figure 4.47 RBMS3 facilitates Slug-induced EMT

MCF7-derived cell lines stably expressed Slug-mCherry vector were plated for 24 hours and treated with TGF-β1. Cells were cultured for 9 days and photographs were taken. Representative images of the cells and fluorescence intensity of mCherry from the same views are shown. Scale bars, 400 μm. Expression level of Slug was confirmed by Western blot (data not shown). Red arrows show the area of cells with mesenchymal-like morphology. The ratio between the longest diameter and the shortest diameter of the cells was determined. Cells for which the ratio was more than 2.1 was regarded as 'spindle-like shaped cells'. Data are shown as mean ± SD from three independent experiments. (*$p < 0.05$, **$p < 0.01$, *** < 0.001)

4.10 RBMS3 and ESRPs play different roles in cell proliferation and migration

To examine the functions of ESRPs in detail and compare with RBMS3, we first performed cell viability assay. Compared to MCF7-Cas9 control cells, loss of ESRP2 demonstrated significantly elevated cell growth, indicative of potent growth suppressive role of ESRP2. Unlike ESRP2, loss of ESRP1 had little effect on MCF7 cell viability. Loss of both ESRPs showed a slight increase of cell growth compared to control, whereas overexpression of RBMS3 induced a significantly increased cell growth in MCF7-ESRP DKO cells. RBMS3 induced little/no increase on cell growth compared with MCF7 or T47D control cells. Loss of ESRP1 caused mild reduction of cell viability compared to T47D-Cas9 control cells, suggesting that ESRP1 has mild growth-promoting function in T47D cells. Conversely, loss of ESRP2 resulted in significant increase in T47D cell viability, confirming the strong growth inhibitory effect of ESRP2 in luminal BC cells. T47D-ESRP-KO cells showed mild reduction of viability compared to T47D-Cas9 control cells, and T47D-ESRP DKO-RBMS3 cells showed significantly increased cell growth compared to T47D-ESRP DKO cells, suggesting that RBMS3 promotes cell growth under depletion of both ESRPs (Figure 4.48). Based on cell migration assay, ESRP1 significantly suppressed MCF7 and T47D cell migration. ESRP2, however, appeared to promote MCF7, but not T47D, cell migration. Loss of both ESRPs significantly promoted T47D, but not MCF7, cell migration compared to control. As expected, RBMS3 induced migration of T47D and MCF7 cells with loss of both ESRPs (Figure 4.49). Scratch assay suggested that cells with RBMS3 or loss of ESRP1 had faster wound closure rate, further confirming the migration suppressive role of ESRP1 in both cell lines. Loss of ESRP2 led to slower closure in MCF7, but not in T47D cells, which is consistent with results of migration assay (Figure 4.50). Compared to Cas9 control cells, loss of ESRP1 greatly facilitated colony formation, whereas loss of ESRP2 inhibited colony formation. Surprisingly, fewer colonies were formed when both ESRPs are lost, this may be due to the contradictory roles of ESRP1 and

ESRP2 on colony formation (Figure 4.51). Mammosphere assay showed that cells with loss of ESRPs, especially ESRP2, formed irregular mammospheres, unlike the rounded ones formed by control cells. In addition, loss of ESRP1 or ESRP2 affected formation of large mammospheres, but not small mammospheres (Figure 4.52-53). When grown in 3D Matrigel, expression of RBMS3 or loss of ESRPs promoted formation of small mammospheres in T47D cells. Conversely, loss of either ESRP1 or ESRP2 showed little effect on mammosphere formation compared with control (Figure 4.54). When cells were cultured in round-bottomed low-attach 96-well plates, we noticed that RBMS3 or loss of ESRP2 appears to promote cell dispersion (Figure 4.55). Collectively, our results indicate that ESRP1, ESRP2 and RBMS3 play distinct roles in BC: While ESRP1 primarily inhibits migration, ESRP2 mainly suppresses proliferation and promotes migration. The opposite functions of ESRP1 and ESRP2 on cell migration may be explained by the requirement of mammary branching morphogenesis during development, possibly controlled by the relative endogenous levels of these two proteins.



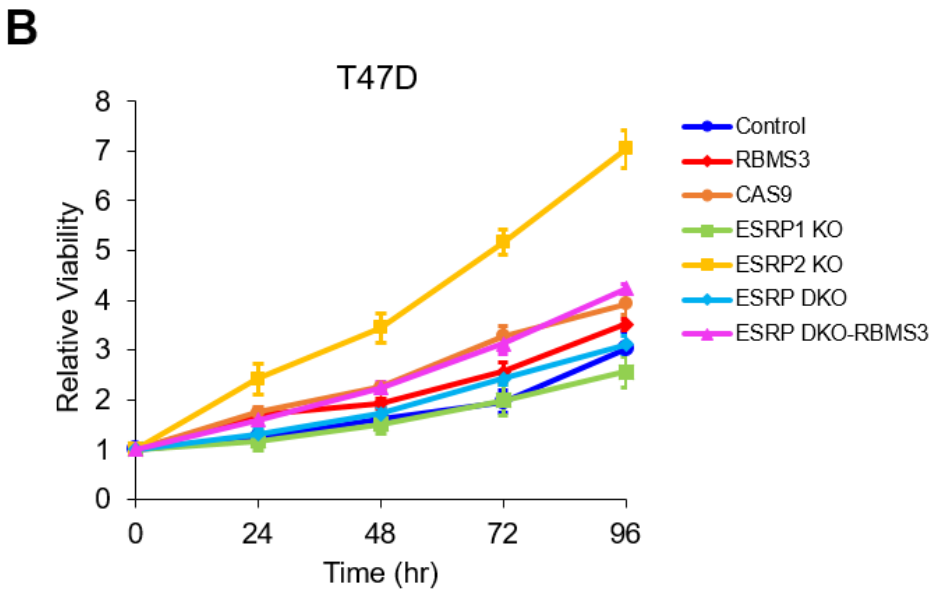
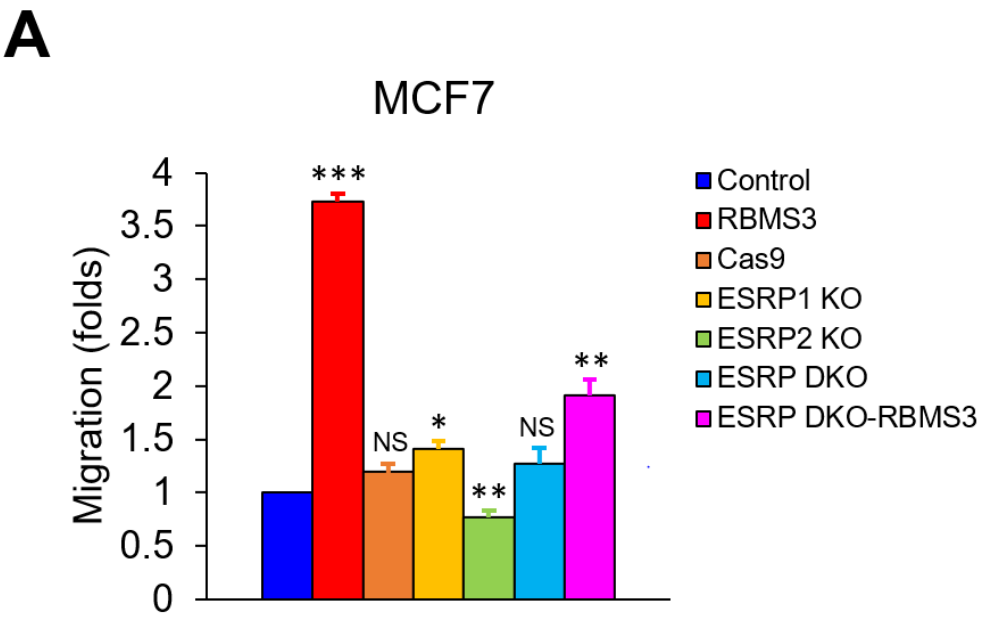


Figure 4.48 ESRP1, ESRP2 and RBMS3 play distinct roles on cell viability
10^3 cells were seeded in 100 µL medium and cultured for indicated time. 50 µL medium was removed and 50 µL Cell-Glo Titer 2.0 reagent added to each well before testing. Plates were shaken 5 minutes and stabilized for another 5 minutes. All values were normalized to MCF7 or T47D cells. Data are shown as mean ± SD from two independent experiments. All values are compared to Cas9 control except that RBMS3-OE is compared to wild-type control. In the MCF7-derived cell lines, $p<0.05$ (ESRP DKO vs. Control), $p<0.01$ (ESRP DKO-RBMS3 vs. Cas9), $p<0.001$ (ESRP2 vs. Cas9), NS (RBMS3 vs. Control, ESRP1 KO vs. Cas9). In the T47D-derived cell lines, $p<0.05$ (ESRP1 KO vs. Cas9, ESRP DKO vs. Cas9), $p<0.001$ (ESRP2 vs. Cas9), NS (RBMS3 vs. Control, ESRP DKO vs. Cas9)



B

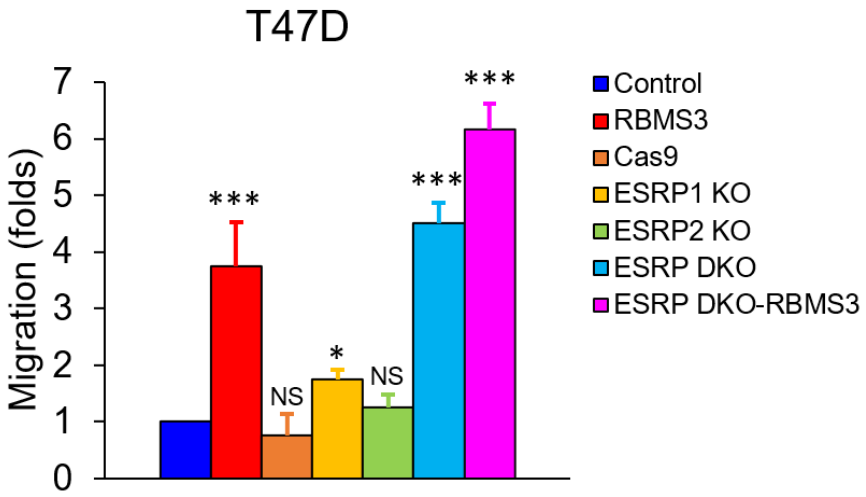


Figure 4.49 ESRP1, ESRP2 and RBMS3 play distinct roles on cell migration
5.0 x10^4 cells were seeded on top of the Matrigel in the upper Boyden chamber and the bottom chamber was filled with culture medium containing EGF (10 ng/mL) as the chemoattractant. After 48 hours, the underside of the Boyden chamber membrane was fixed, stained and counted. Data are shown as mean ± SD from three independent experiments. All values are compared to wild-type control. (*p < 0.05, **p < 0.01, *** < 0.001)

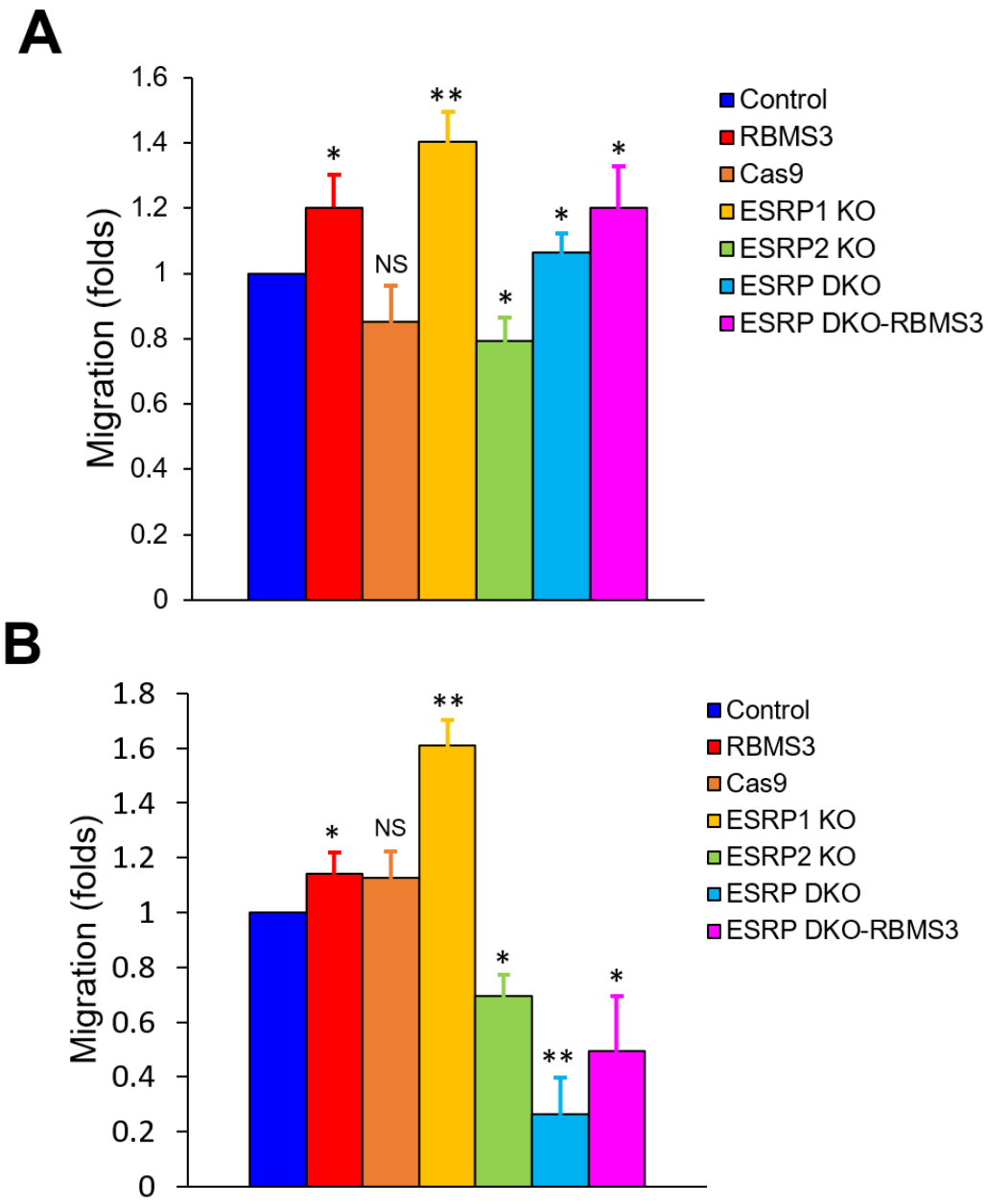


Figure 4.50 ESRP1, ESRP2 and RBMS3 play distinct roles on wound healing

MCF7-derived (A) and T47D-derived (B) cells were grown to around 90% confluency and serum-starved overnight. A scratch ('wound') was inflicted to the cell layer. Wound closure was photographed at 0 and 24 hour. Cell migration rate was calculated upon the difference of gap width at 24 hour compared to 0 hour depending on the speed of wound closure. Data are shown as mean ± SD from two independent experiments. All values are compared to wild-type control.

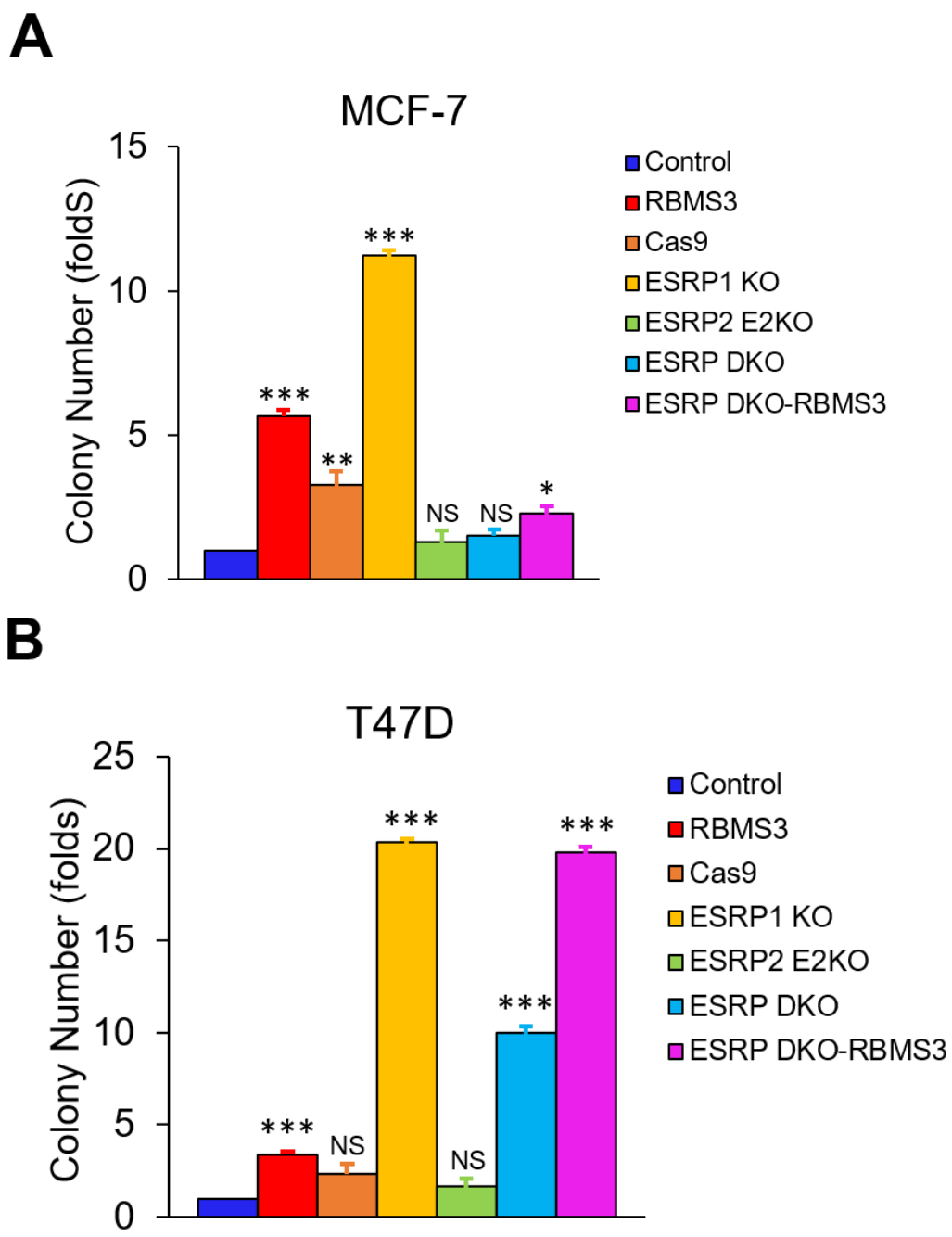


Figure 4.51 ESRP1, ESRP2 and RBMS3 have distinct roles on colony formation

10^3 cells were seeded in the upper layer of the soft agar plate and grown for two weeks. Colonies were stained with crystal violet and counted under four random views. Data are shown as mean ± SD from three independent experiments. (**p < 0.01)

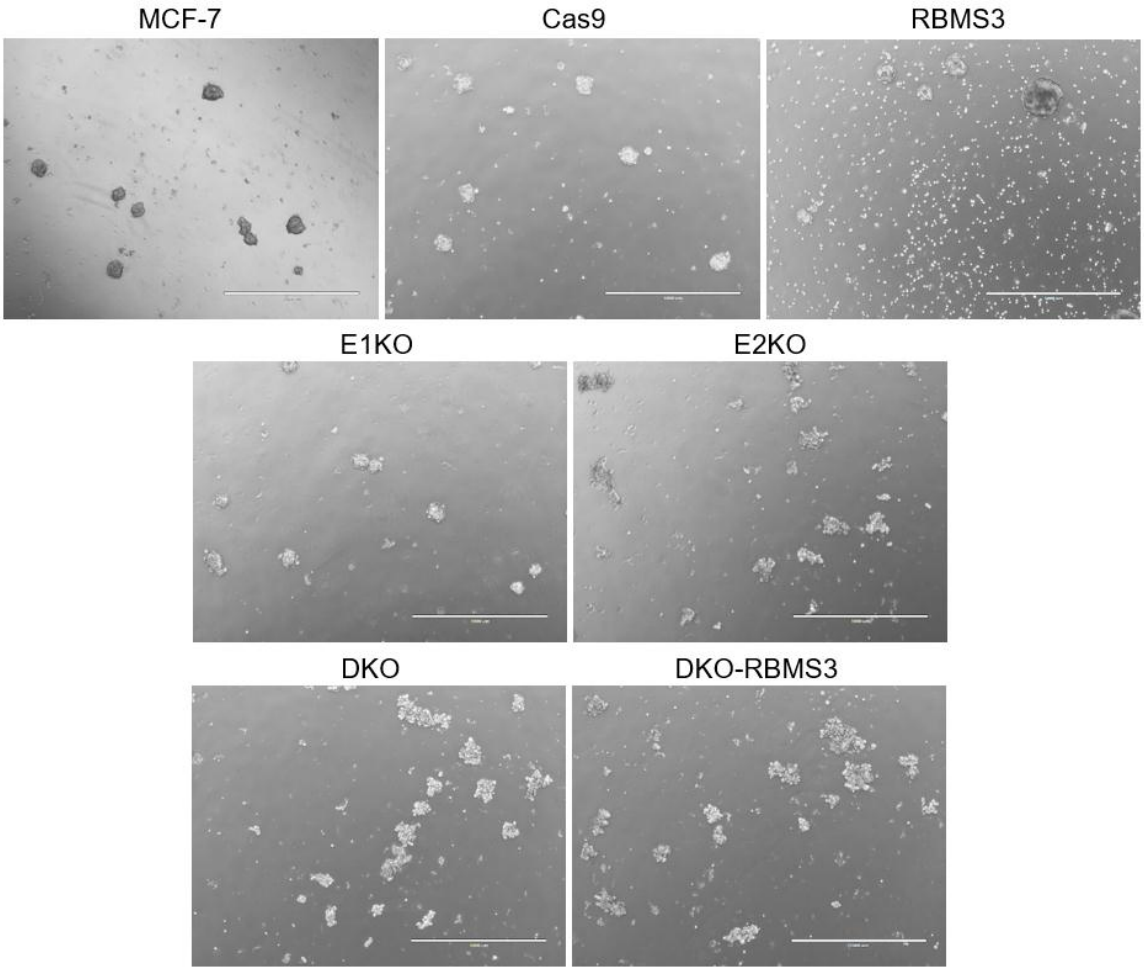


Figure 4.52 ESRP1, ESRP2 and RBMS3 have distinct effects on mammosphere formation

Tumorsphere formation was assessed in MCF7 cell with or without loss of ESRPs and in the presence or absence of RBMS3. 10^5 cells were seeded in 6-well plate using commercial mammosphere-forming media and cultured for two weeks. Representative images of seven MCF7-derived cell lines are shown. Scale bars, 1000 μm.

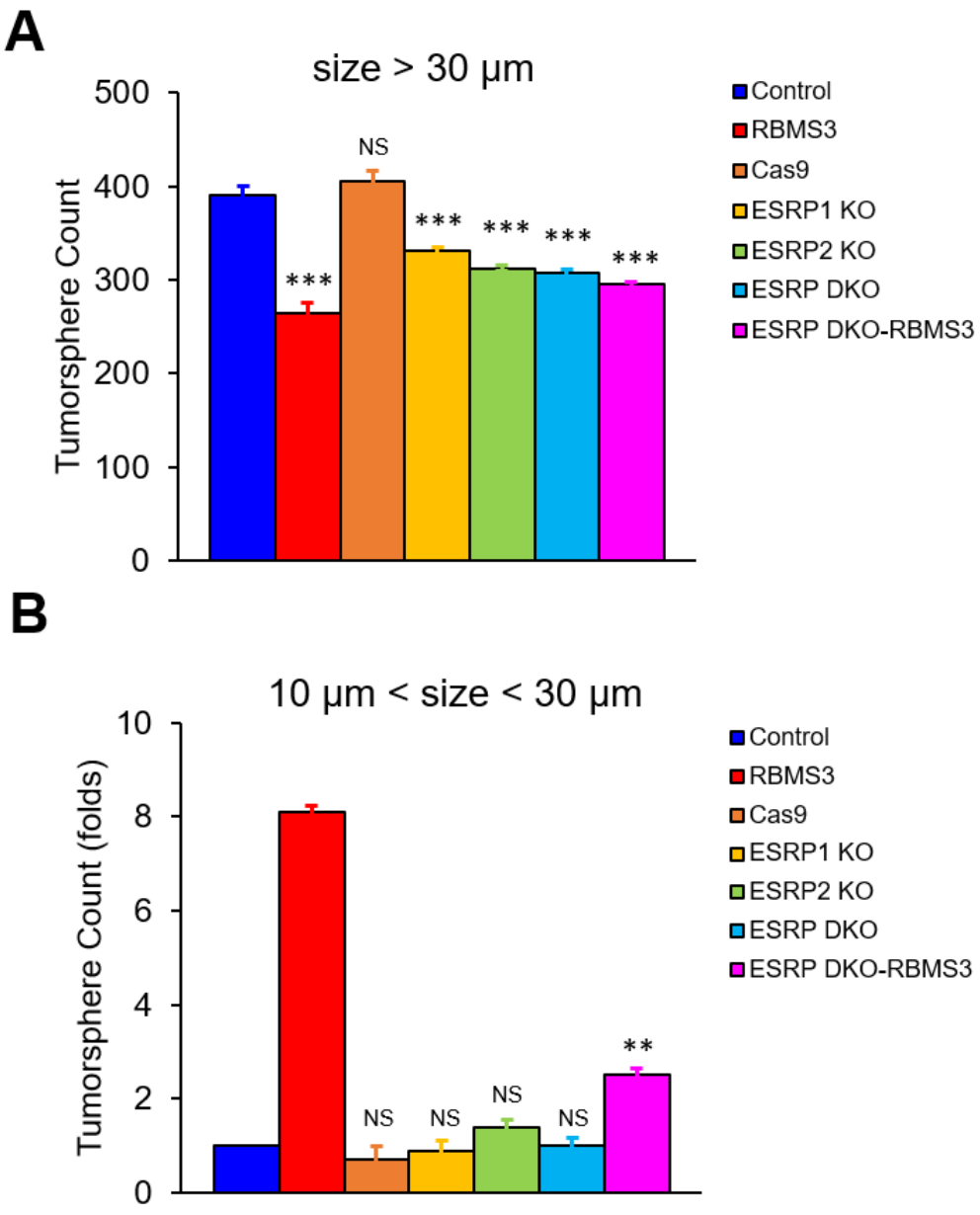


Figure 4.53 RBMS3 or loss of ESRPs inhibits mammosphere formation
Graphic representation of tumorsphere counts based on tumorsphere size (30 μm) in MCF7-derived cell lines from tumorsphere formation assay. Data are shown as mean ± SD from twenty independent views.

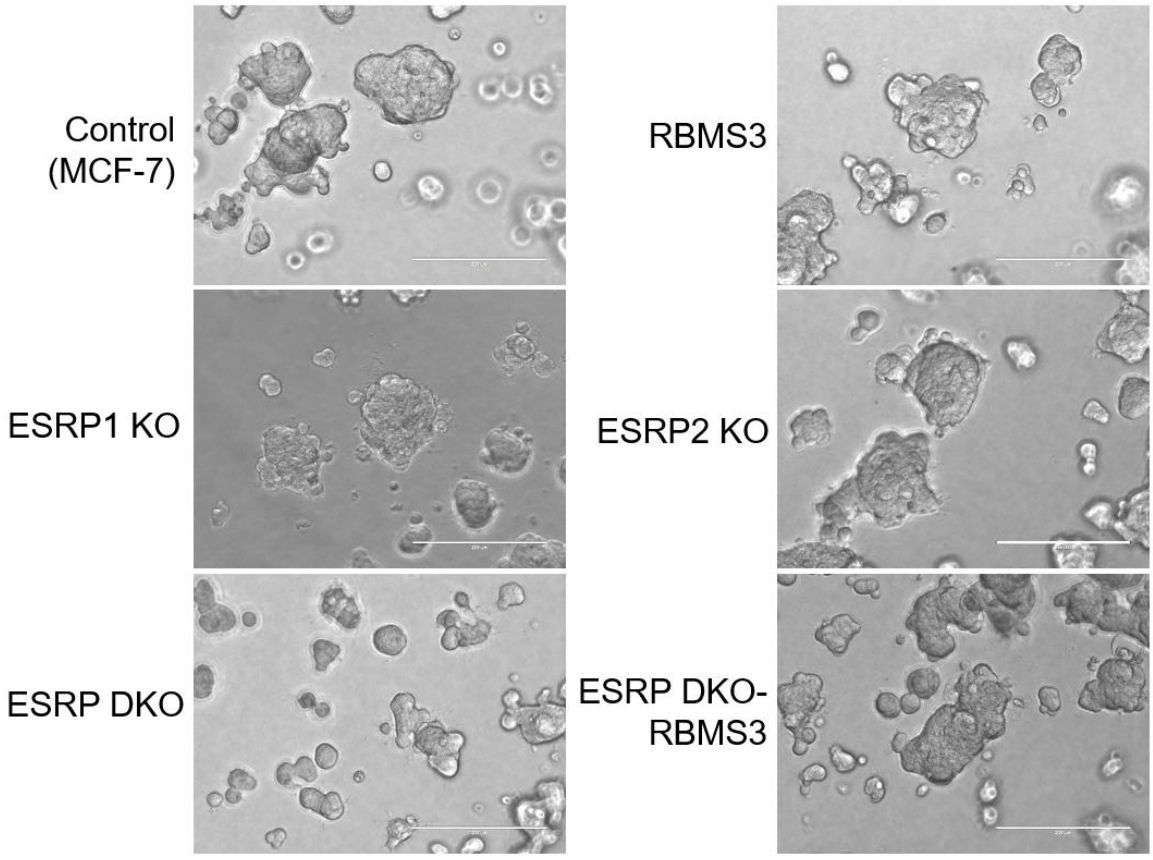


Figure 4.54 RBMS3, or loss of ESRPs, promotes formation of small tumorsphere clusters

Representative images of 3D on-top Matrigel culture in MCF7-derived cell lines.1.5x 10^4 cells were seeded on a layer of Matrigel and 10% Matrigel was covered on top. Cells were culture and photographed at day 4.

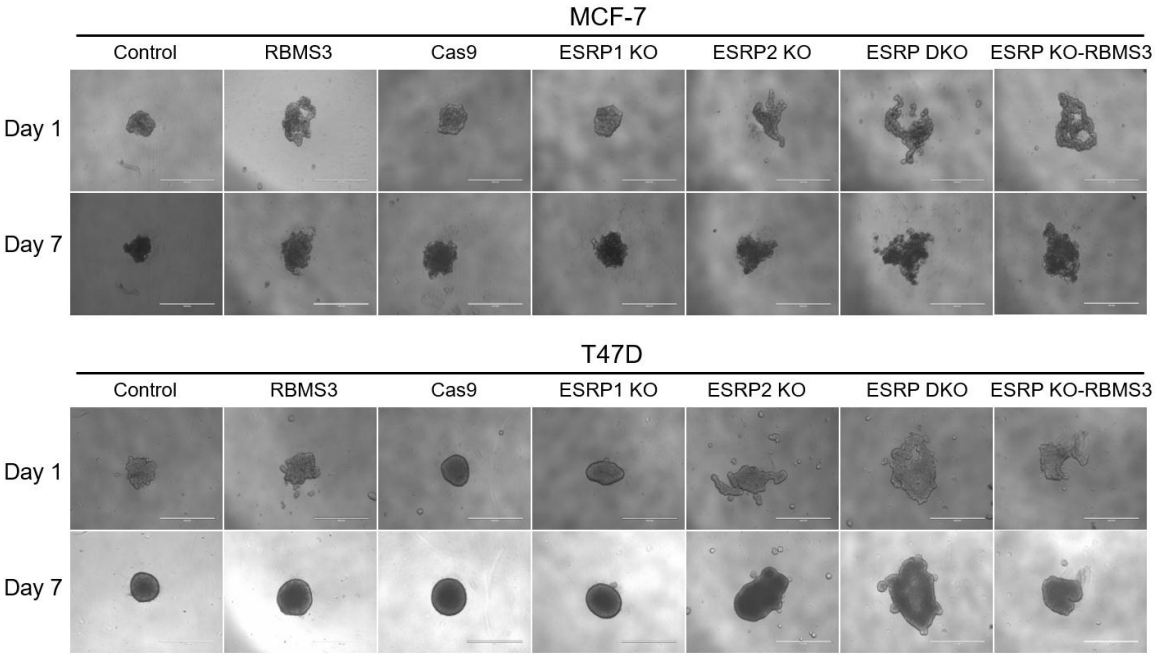


Figure 4.55 Representative images of adhesion-independent growth of luminal BC cells with or without ESRPs and/or RBMS3

Adhesion-independent growth of 14 luminal BC-derived clones were examined. Briefly, $0.8x10^3$ cells were seeded in 96-well round bottom ultralow-attachment plates. Cells were cultured for up to one week and photographs were taken daily. Shown are representative images of day 1 and day 7. All cells are seeded in triplicates.

4.11 Identification of Alternative Splicing Events Regulated by ESRPs and RBMS3

While much is known about transcriptional regulation of EMT, alternative splicing of genes and their contributions to the morphological conversion accompanying EMT have not been comprehensively investigated. Although ESRP1 and ESRP2 have been shown to inhibit AS events during EMT, it remains unclear whether RBMS3 also alter splicing of genes and if these three RBPs share common targets that are directly involved in EMT. To dissect the mechanism by which ESRP1, ESRP2 and RBMS3 regulate EMT, we used the established cell culture model and RNA-seq analysis to assess changes in transcriptional programs. Briefly, as mentioned above, we knocked out each or both of the ESRP1 and ESRP2, or overexpressed RBMS3 in T47D cells. A total of seven clones (T47D, T47D-Cas9, T47D-RBMS3, T47D-ESRP1 KO, T47D-ESRP2 KO, T47D-ESRP DKO and T47D-ESRP DKO-RBMS3) were sent for RNA-seq. To deduce regulon of the three RBPs from

our dataset, we paired matched the expression results of each clone to generate lists of DEGs (marked as A, B, C, D, E, and F) potentially regulated by the corresponding RBP (Figure 4.56). By cross-checking each list, information of potential regulon was obtained. Depletion of ESRPs affected expression of 4705 transcripts, 49.3% of which were downregulated, whereas overexpression of RBMS3 affected a smaller set of 1307 transcripts, 61% of which were upregulated (Figure 4.57A). Interestingly, by comparing lists B and C, significantly less transcripts were affected by RBMS3 when ESRPs were already depleted, revealing similar regulatory functions for these RBPs. Comparison of lists E and F showed that the number of genes regulated by ESRPs and RBMS3 is more in 'EMT' than in 'MET' process (Figure 4.57B). Among the 71 and 2096 consistently regulated transcripts, only 14 were potentially shared by ESRP(s) and RBMS3 (Figure 4.57C). A heatmap of the top 60 DEGs is shown for RBMS3 and ESRP depletion lists (Figure 4.57D).

We also performed AS analysis in RNA-seq data using R software to individually quantify and analyze differences in AS events. The number of top AS events were listed and compared. Among all AS events, 39 events were shared by ESRP1 and ESRP2. Another 39 events were shared by ESRP1 and RBMS3; and 53 were shared by ESRP2 and RBMS3. Interestingly, one event was found to be commonly regulated by all three RBPs (Figure 4.58A). Next, we examined five common types of alternative isoform expression events, each capable of producing multiple mRNA isoforms from a gene through AS. These AS events included 3' alternative splicing (A3SS), 5' alternative splicing (A5SS), exon skipping (SE), retained intron (RI) and mutually exclusive exon (MXE). A comprehensive set of 9808 events of these five types was derived from the analysis in T47D-RBMS3 versus control cells. 9461 and 9874 events were observed in T47D-ESRP1 KO and T47D-ESRP2 KO versus control cells, respectively. These events were then ranked to generate top splicing lists based on reads. As observed for transcriptional targets, the overlap between alternatively spliced exons regulated by these RBPs was moderate. In

T47D-RBMS3 versus control T47D cells, MXE event occurs most frequently in the top splicing list. A3SS and A5SS events occur at similar moderate levels, whereas RI event appears least frequently (Figure 4.58B).

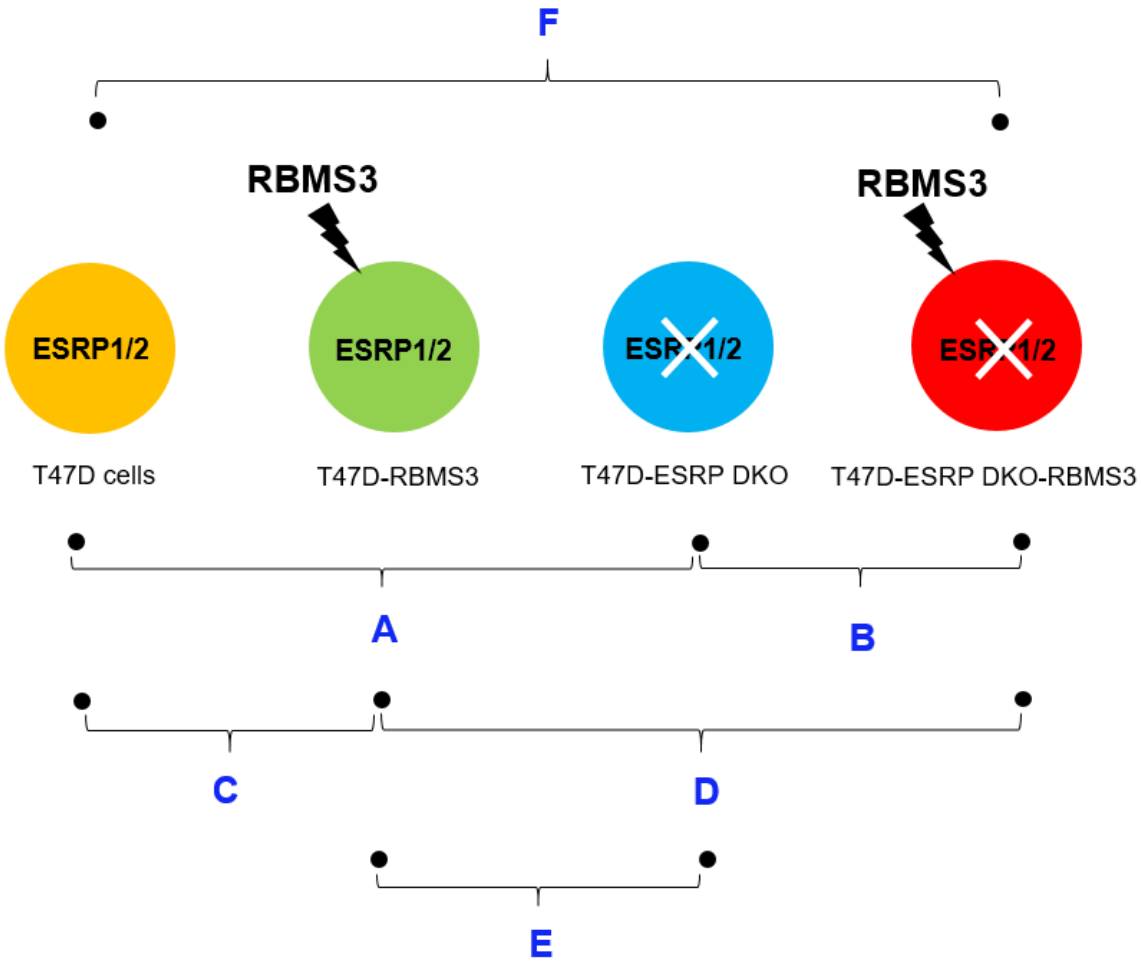
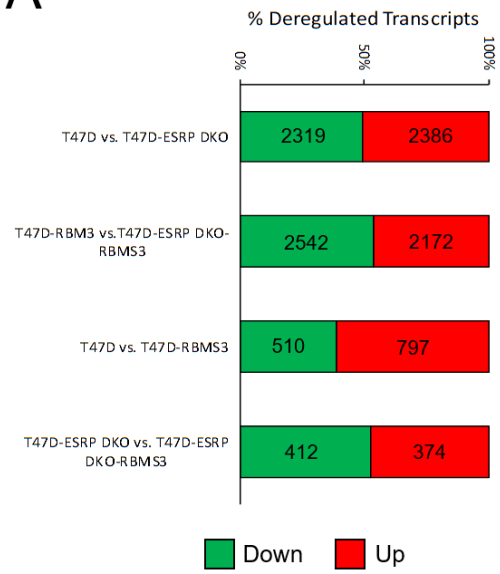


Figure 4.56 Schematic of analytic workflow of RNA-seq dataset
Expression sets of these four cell lines were pair-matched and compared by FC, generating lists of DEGs marked as group A, B, C, D, E and F. FC greater than 0.5 is considered as differentially expressed. By cross-checking each list, information on the potential regulon of each RBP will be inferred.

A

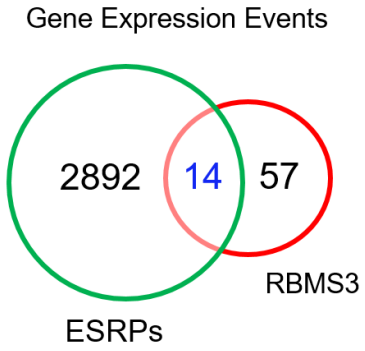


B

List	Deregulated Genes	Regulated By
A	4705	ESRP1/2
B	786	RBMS3
C	1307	RBMS3
D	4714	ESRP1/2
E	4717	ESRP1/2 and RBMS3
F	4543	ESRP1/2 and RBMS3

C

Lists	Genes Present in Both Lists	Genes Consistently Upregulated/Downregulated
B+C	199	71
A+D	3083	2096



D

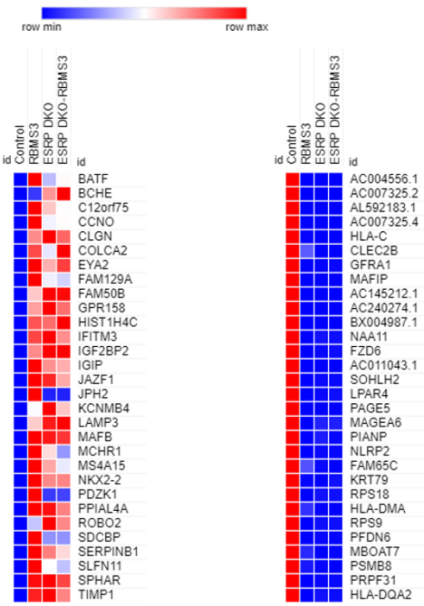


Figure 4.57 Expression Regulation by ESRP1, ESRP2 and RBMS3
(A) Number and percentage of transcripts up- or downregulated more than 1.5 fold upon depletion/overexpression of the indicated RBPs.

(B) Number of DEGs (>1.5 fold) present in each list.

(C) Number DEGs consistently present or consistently regulated in both list B and C, or A and D. Circles represent the number of DEGs regulated by the indicated RBPs. Overlap between gene expression changes is shown in blue.

(D) Heatmap for the top 60 DEGs in RNA-seq dataset.

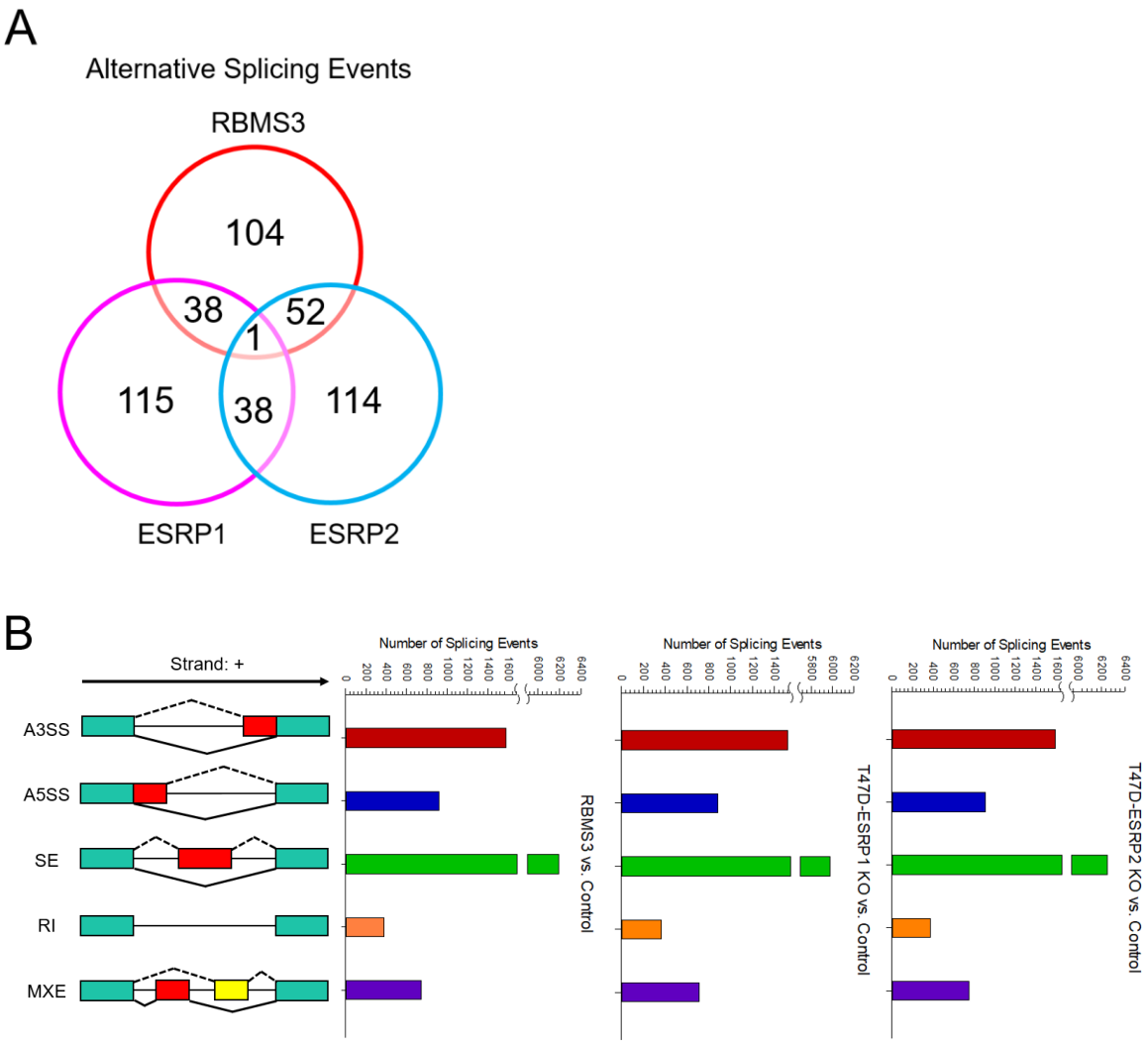


Figure 4.58 Splicing Regulation by ESRP1, ESRP2 and RBMS3
(A) Overlap between alternative splicing changes observed upon RBP depletion or overexpression.

(B) Number of splicing changes upon depletion of the RBPs across categories of alternative splicing events. A3SS = 3' alternative splice site, A5SS = 5' alternative splice site, SE = exon skipping, RI = retained intron, MXE = mutually exclusive exon

4.12 Proposed Model of ESRPs and RBMS3 in BC

Normal mammary cells express potent epithelial regulators, such as ESRPs, to govern epithelial identity and suppress EMT. In normal ductal cells and luminal BC cells, TGF-βr signaling remains at minimal/undetected level. Upon stimulation by external TGF-

β, cells respond through an intricate balance of Smad-dependent or independent signaling, eventually driving the tumor-supressing arm of TGF-βr signaling. During advanced BC stages, chronic activation of high levels of TGF-βr signaling disables/inactivates the tumor-suppresing arm, which shifts the equilibrium to the tumor-promoting arm and leads to activation of oncogenes like *RBMS3*. RBMS3 upregulates Smad proteins and further enhances TGF-βr signaling, which induces expression of EMT master regulators, eventually leading to metastatic behaviors of cancer cells, characterized by loss of epithelial polarity and conversion into an elongated, migratory and invasive phenotype (Figure 4.59) [55].

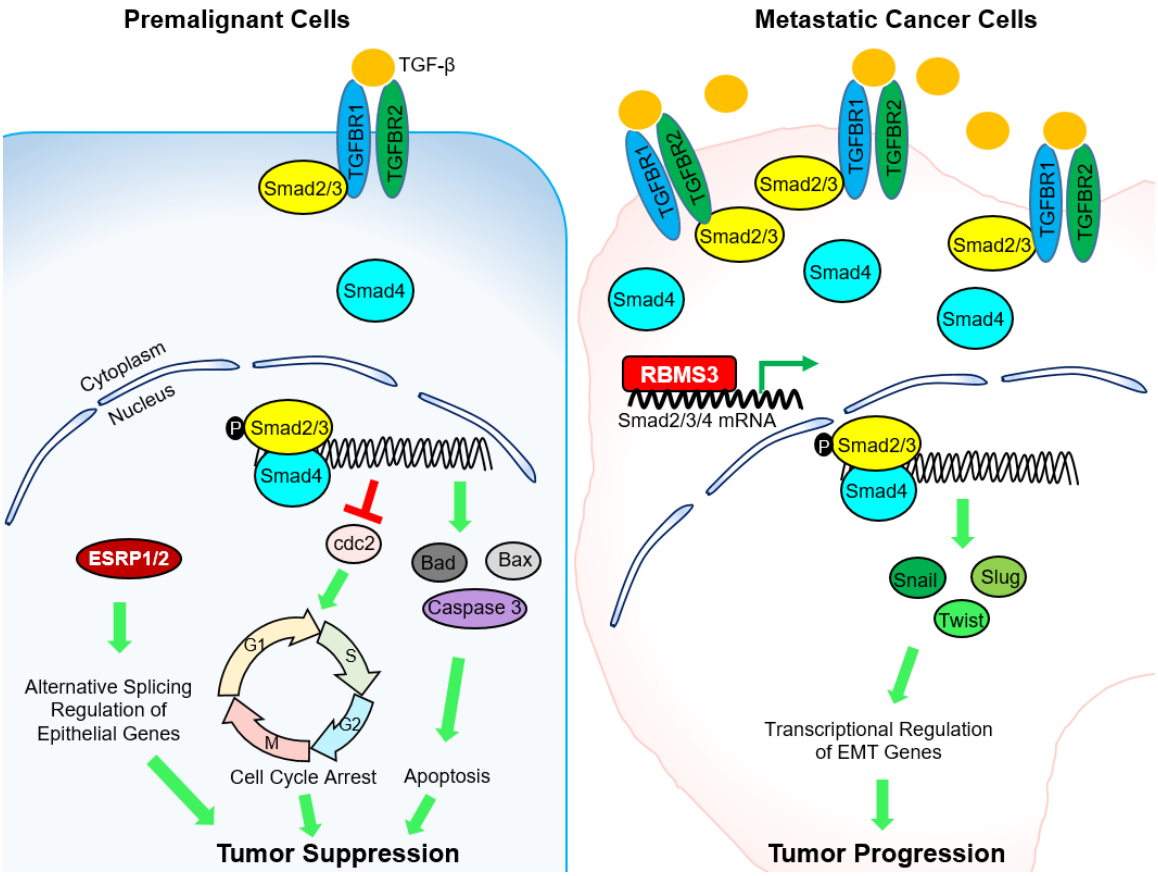


Figure 4.59 Proposed model on RBMS3 and ESRPs

Mammary ductal cells possess intrinsic apical-basal epithelial polarity maintained by several types of cellular junctions. TGF-β signaling is initiated by ligand-induced oligomerization of serine/threonine receptor kinases and phosphorylation of the cytoplasmic molecules Smad2/3, resulting in their interactions with transducer Smad4 and translocation into the nucleus. Activated Smads regulate diverse biological effects by partnering with other transcription factors, leading to cell-state specific modulation of transcription. Specifically, in premalignant cells, cell cycle arrest and apoptosis are two common pathways that are regulated by TGF-β signaling to suppress tumor formation. On the contrary, during later stage of tumorigenesis, the tumor suppressing arm of TGF-β is disabled, shifting to tumor-promoting arm of TGF-β signaling, leading to activation of oncogenes like *RBMS3*. RBMS3 upregulates TGF-β-regulated EMT master regulators (Snail, Slug and Twist, etc) and promote tumor progression by transcriptional control of EMT genes, eventually resulting in tumor progression. (Orange shape indicates ligand; Blue and purple hooked shapes indicate serine/threonine receptor kinases; Beige egged shape with black outline indicates nucleus; Black turning arrow indicate 'activation of transcription'; Double wave indicates DNA; Single wave indicates RNA)

CHAPTER 5. DISCUSSION

For the past decade, increasing attention has been drawn to the post-transcriptional network of EMT, which involves alternative splicing, noncoding RNA and mRNA stability, etc. RBMS3 is a possible post-transcriptional regulator that popped out from our microarray dataset. So far the function of RBMS3 in tumorigenesis has been seemingly controversial. Despite being downregulated in TCGA overall breast tumor patients compared to normal counterparts, RBMS3 shows increased expression among invasive BLBC cells. Hence, it may be difficult to interpret the actual function of RBMS3 solely based on TCGA data due to heterogeneous nature of breast tumors. Intriguingly, Oncomine dataset also showed that *RBMS3* is highly expressed in the immune cells, which indicated *RBMS3* could potentially play a role in the immune system (data not shown). Another interesting conundrum is that, while the expression of both ESRPs are low in normal epithelium, they are upregulated in luminal BC cells and downregulated in invasive fronts.

In an attempt to study RBMS3 and ESRPs in detail, we took the advantage of several BLBC cell lines and luminal BC cell lines. These two subtypes of BC cells are commonly used to study the EMT process as they nicely represent the epithelial state and mesenchymal state, respectively. As luminal BC cells and basal-like BC cells arise from genetically-different backgrounds, we were not unexpected to see inconsistencies on the viability assay between MCF7 cell and BT-549 cells. Failure to elicit effect on proliferation in MCF7 cells does not imply that RBMS3 cannot not involved in cell proliferation, as its effect could neutralized or mitigated by other potent factors in the MCF7 cells. Surprisingly, unlike previous studies that revealed largely redundant roles for ESRP1/2 on cell motility in several cancers [153-155], we observed differential functions of ESRP1/2 on luminal BC cell migration and proliferation. While ESRP1/2 are essential to suppress EMT, loss of ESRPs is not sufficient to drive mesenchymal

phenotype in luminal BC cells, suggesting that other core molecules may be involved in the EMT process. Our functional assays revealed that RBMS3 has important roles on cell migration, invasion and CSC self-renewal. Specifically, CSCs are thought to be responsible for drug resistance and disease relapse, and the rationale behind mammosphere assay is that only CSCs are able to form individual mammospheres in the specialized medium under suspension condition. Compared to the control cells that formed regular-sized and sphere-shaped mamospheres, RBMS3 overexpression led to significantly increased numbers of small and irregular-shaped clusters which tend to be quite loose and unstable, reminiscent of what is observed in the MCF7 cells with knockdown of E-cadherin. Indeed, formation of small and loose cell cluster appears to be more favorable for cancer cell dissemination in the circulation as it may increase the efficiency of spreading and the chances of forming metastasis.

To study whether RBMS3 serves as a central player in the EMT process, we induced RBMS3 expression in T47D and MCF7 luminal BC cells as well as in those with loss of both ESRPs. T47D and MCF7 cells are widely studied in breast cancer mechanisms. While both being estrogen receptor-positive, they have intrinsically distinct genetic background and molecular profiles as previous studies from 2D gel and mass spectrometry have revealed that several proteins involved in cell growth stimulation and anti-apoptotic mechanisms are more strongly expressed in T47D than MCF7 cells. Additionally, MCF7 cells are reported *TP53* wildtype, while T47D cells are *TP53* mutant. These reasons may be at least partially responsible for the inconsistencies between MCF7- and T47D-derived cell lines observed in our functional assays. Specifically, since T47D-ESRP DKO and T47D-ESRP DKO-RBMS3 cells grew in clusters and attached loosely to culture dish, it may be possible that a small portion of the cells were lost at medium removal step during cell viability assay, affecting the final measurement and leading to lower values. Unfortunately, we failed to observe notable morphological or molecular changes to mesenchymal phenotype. However,

upon overexpression of EMT transcription factor Slug, we observed that T47D-ESRP DKO-RBMS3 cells transformed into a typical mesenchymal morphology, characterized by spindle-like shape, at a much faster rate than that of the T47D-ESRP DKO cells, indicating that T47D-ESRP DKO-RBMS3 cells reside in a place that is closer to the mesenchymal state than the T47D-ESRP DKO cells. These data also inform us that other essential factors may also be engaged in the EMT process. To elucidate this, co-immunoprecipitation of RBMS3 followed by liquid chromatography-mass spectrometry (LC-MS) may be performed to identify potential RBMS3-interacting partners. Alternatively, we may perform stable isotope labeling by essential nutrients in cell culture (SILEC) for LC-MS analysis. Furthermore, loss of both ESRPs conquered the apoptotic effect induced by long term expression of EMT transcription factors, suggesting that ESRPs may be responsible in regulating apoptotic pathways that are suppressed in mesenchymal cells.

To date, the mechanistic role of RBMS3 in human development has not been established. The one and only study for in vivo function of RBMS3 was described in the zebrafish model, where RBMS3 is transiently expressed in the cytoplasm of migrating cranial neural crest (CNC) in the pharyngeal arches during cartilage differentiation. RBMS3 spatiotemporally controls the timing and duration of the TGF-βr signaling within CNC cells through transcript stability of TGF-βr pathway components, which influences proliferation and initial differentiation of prechondrogenic CNC. TGF-βr pathway promotes Sox9-dependent transcriptional activity on the *a1 (II) collagen* gene (*COL2A1*) enhancer to induce chondrogenesis [156]. That being said, a certain threshold of TGF- βr signaling is likely to be required to drive prechondrogeneic cells down the cartilage lineage. Once become committed to chondrogenesis, *Rbms3* expression is shut down in these cells [136]. For years TGF-β, a potent pleiotropic cytokine, has been known as an inducer of cell arrest in benign cells and early-stage tumors, but also a contributor of breast cancer cellular heterogeneity and metastasis. This mysterious

phenomenon or conundrum is considered as the 'TGF-β' paradox. Unfortunately, due to the dual nature of TGF-β in BC patients, it is unlikely to be utilized as a tumor marker to distinguish patients with high risk of metastasis [157]. To date, regulators involved in TGF-β signaling during tumor progression and the underlying mechanisms still remain to be clarified. Despite of difficulty, increasing evidence has revealed that the diversity of TGF-β response is determined by combinatorial usage of core pathway components [158]. Therefore, unraveling the activities of these central players is crucial to understand the execution of TGF-β-mediated EMT. It is known that expression of E-cadherin is suppressed by Snail [47], which is shown in this study to be induced by RBMS3-mediated TGF-β/Smad signaling.

We based our study on the finding in the zebrafish model. In our studies, we verified the link between TGF-β and RBMS3 in breast cancer cells and provided detailed mechanisms for RBMS3-mediated activation of TGF-β/Smad signaling. Our findings on RBMS3 have at least four broader impacts: 1) It furthers our understanding of TGF-β pathway; 2) It enables the possibility of using RBMS3, instead of TGF-β, as a potential tumor marker to distinguish patients with higher probability of metastasis; 3) It facilitates the discovery of other essential mRNA regulators important for EMT; 4) Understanding the importance of RBMS3 help us to further study the role of RBMS3 in human development and other biological processes.

In bone-metastatic cancer cells, connective tissue growth factor (CTGF) is highly expressed and secreted [159]. CTGF is known to be expressed at low levels in luminal BC cells and found to mediate TGG-β-induced apoptosis in MCF7 cells. Interestingly, we found significant downregulation of CTGF in luminal BC cells with expression of RBMS3, suggesting that even though RBMS3 may conquer TGF-β-mediated apoptosis and cytostasis, its expression alone failed to elicit TGF-β-mediated metastasis in luminal BC cells. We also noticed that depletion of RBMS3 downregulated *Smad2* and *Smad3* but not *Smad4*. This suggest that RBMS3 tightly controls the total pool of *Smad2* and

Smad3 through regulating mRNA stability. The inconsistency between qPCR and Act D stability assay, reflected by unaffected level of *Smad4* mRNA and decreased *Smad4* mRNA stability in BT-549 RBMS3-KO compared to control cells, suggest that *Smad* may also be regulated through other post-transcriptional or translational mechanisms. Interesting, when we induce BT549 control and RBMS3-KO cells with TGF-β1, we observed downregulation of Slug and Twist and upregulation of Snail in BT549 cells that is absent in the RBMS3-KO cells, suggesting a possible feedback loop for the cells to maintain the equilibrium of EMT genes.

Using Kaplan-Meier Plotter, we discovered a correlation between RBMS3 expression and prognosis in BC. The prognosis of advance stage BC patients with high expression RBMS3 was poor compared with those with low expression of RBMS3. Here we demonstrated that RBMS3 is closely associated with BC malignancy, suggesting RBMS3 may serve as a novel prognostic marker and therapeutic target for malignant BC. Beyond our study is whether RBMS3 exert the same tumor-promoting roles in vivo. As for future directions, in vivo studies should be performed to find out whether RBMS3 is critical for tumorigenicity of TNBC and whether RBMS3 inhibitors may suppress the tumorigenicity by inhibiting TGF-β pathway.

From our RNA-seq analysis, we discovered several expression events that are potentially co-regulated by ESRP1, ESRP2 and RBMS3. However, whether these DEGs are direct targets or consequence of 'butterfly effect' is still unknown. Additionally, we also identified several promising AS events. These targets should be validated in order to identify critical ones that are indispensable and consistently changed during the EMT process. The study of these DEGs will greatly enhance our understanding of the post-transcriptional network and landscape of EMT and metastasis. Additionally, we identified one AS event co-regulated by ESRPs and RBMS3. This AS event was unpublished and leads to the production of two isoforms of the gene product, which is a membrane protein that has at least 2-3 isoforms. One isoform was reported to be closely

related to mammary tumors. Therefore, it will be interesting to study whether ESRPs and RBMS3 truly regulate the splicing of this gene and how this process contributes to EMT.

The characterization of RBMS3 expression in BC during my dissertation research, including the phenotypic and mechanistic data, answered several questions about the tumor promoter role of RBMS3. However these data have also raised several more important pertinent questions. As for further directions, firstly, it may be necessary to study the biological function of RBMS3 in greater detail. While RBMS3 has been shown to reduce cell contact and E-cadherin, it will be interesting to study how RBMS3 regulates polarity. In addition, since RBMS3 belongs to the MSSP family and other members are found to be involved in gene transcription, cell cycle and apoptosis, we anticipate that the role of RBMS3 on TGF-β-mediated cytostasis and apoptosis is mediated through transcriptional regulation of *c-myc*. Secondly, the study of splicing mechanisms of ESRPs and RBMS3 will provide detailed information on the post-transcriptional landscape of EMT. Thirdly, according to TCGA data on Oncomine website, RBMS3 shows a differential expression pattern in immune cells of breast carcinoma. Fourthly, a published microarray indicated that knockdown of ER induced expression of *RBMS3*, suggesting a potential positive correlation on the hormonal control of *RBMS3*. Oppositely, we expect that RBMS3 expression will lead to reduction of ER, leading to the ER-negative phenotype. Fifthly, due to the predominant presence of RBMS3 in the cytoplasm, it might be involved in other aspects of the RNA metabolism, such as RNA localization, mRNA decay, translational regulation, etc. Lastly, the role of RBMS3 in cancer has been controversial. A number of studies have shown as RBMS3 as a tumor suppressor in pancreatic cancer, epithelial ovarian cancer and epithelial esophageal carcinoma. However, results from our study have provided opposite lines of evidence for RBMS3, resulting in the conflict within the field.

Therefore, it will be necessary to address the *in vivo* function of RBMS3 in BC by IHC analysis and mouse studies.

Beyond this study are some bigger questions that are waiting for answers. Firstly, it is still unclear whether cells at the invasive front are predetermined to serve a leader role or are simply induced to a leader phenotype due to their close interactions with the environment. Is RBMS3 expressed only in the leader cells? When do cells begin to express RBMS3? Is RBMS3 silenced during mesenchymal-epithelial transition (MET) process after seeding metastasis? How is RBMS3 regulated by the spatiotemporal control of cell signals? In order to find out the answers, it may be necessary to investigate the micro-environmental context in greater details. Secondly, the on-off switch that drives the expression of RBMS3, which subsequently triggers TGF-β to work as a tumor promoter during metastasis, is still unknown. Lastly, the mechanism underlying the transition of collective movement into widespread dissemination in the metastatic cascade also remains an open question.

In summary, the study within this dissertation has uncovered a pivotal role of RBMS3 in maintaining mesenchymal identity and promoting EMT in BC. This study allows us to explore the potential of using RBMS3 as a classification marker to stratify high grade tumors with high levels of TGF-β signaling and a diagnostic marker for developing TNBC treatment strategies. It also provides directions on the identification of RBMS3-regulated DEGs or splicing patterns that may also serve as diagnostic markers. These studies will greatly enhance our understanding of the post-transcriptional regulation of EMT and breast cancer progression.

Table 5 EMT and apoptosis status in T47D-derived cells with overexpression of Slug/Snail/Twist or TGF-β treatment (10 ng/mL) at day 9 (post transfection or treatment).

Cell Lines	Slug		Snail		Twist		TGF-β	
	EMT	Apoptosis	EMT	Apoptosis	EMT	Apoptosis	EMT	Apoptosis
Control	-	+++	-	+++	-	+	-	+++
RBMS3	-	+	-	+	-	-	-	+
ESRP1 KO	-	+	-	+	-	+	-	-
ESRP2 KO	-	+	-	+	-	-	-	-
ESRP DKO	-	-	-	-	-	-	-	-
ESRP DKO-RBMS3	++	-	++	-	-	-	+	-

*The number of '+' represents the degree or extent of apoptosis, which is measured by the number of viable cells. Note that we summarize 'EMT' status only based on typical spindle-like appearance instead of other characteristics or phenotypes.

APPENDIX

List of Acronyms and Abbreviations

Abbreviation	Acronym
ANK	Ankyrin repeats
ARE	AT-rich element
AUF1	AU-rich element RNA binding protein 1
BLBC	Basal-like breast cancer
CCLE	Cancer cell line encyclopedia
Co-Smad	Co-regulatory Smad
CSC	Cancer stem cell
CTBP	C-terminal-binding protein
CTGF	Connective tissue growth factor
DEG	Differentially expressed genes
DMEM	Dulbecco's Modified Eagle Medium
E-cad	E-cadherin
ELAV	Embryonic lethal abnormal vision
EMT	Epithelial-Mesenchymal Transition
ER	Estrogen receptor
ESRP	Epithelial splicing regulatory protein
FACS	Fluorescent Activated Cell Sorting
FBS	Fetal Bovine Serum
FC	Fold change
FGFR	Fibroblast growth factor receptor
GO	Gene ontology
HEPES	Hydroxyethyl piperazineethanesulfonic acid
HER2	Human epidermal receptor 2
HMGA2	High mobility group AT-Hook 2
hTERT	Human telomerase reverse transcriptase
IHC	immunohistochemistry
JAK	Janus kinase
JNK	c-Jun N-terminal kinase
KC	Chemokine ligand 1
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	Knockout (of a gene)
LC-MS	liquid chromatography-mass spectrometry
MAPK	Mitogen-activated protein kinase

MAPKKK	Mitogen-activated protein kinase kinase kinase
MaSC	Mammary stem cell
MCODE	Molecular Complex Detection
MET	Mesenchymal-Epithelial Transition
N/A	Not applicable
NICD	Notch intracellular fragment
NLS	Nuclear localization signal
OE	Overexpression
ORF	Open reading frame
PARP	poly(ADP-ribose) polymerase
PD-L1	Programmed death ligand 1
PEST	Proline (P), glutamic acid (E), serine (S) and threonine (T)
PKC	Protein kinase C
PR	Progesterone receptor
qPCR	Quantitative Real-time Polymerase Chain Reaction
RAM	RBPJ-κ association module
RBD	RNA binding domain
RBMS3	RNA binding motif single-stranded protein 3
RBP	RNA binding protein
RBPJ	Recombination Signal Binding Protein For Immunoglobulin Kappa J Region
RRM	RNA recognition motif
SILEC	Stable Isotope Labeling by Essential nutrients in Cell culture
SMAD	Homologies to the *Caenorhabditis elegans* SMA ('small' worm phenotype) and Drosophila MAD ('Mother Against Decapentaplegic') family of genes
R-Smad	Regulatory Smad
RTK	Receptor tyrosine kinase
Ser-Thr	Serine and threonine
SFRP1	Secreted frizzled related protein 1
snRNP	Small nuclear ribonucleoprotein
TAD	Transcriptional activation domain
TAK1	Transforming growth factor β-activated kinase 1
TIL	Tumor-infiltrating lymphocytes
TβR	TGF-β receptor
TGF-β	Transforming growth factor beta
TNBC	Triple-negative breast cancer
TPM	Transcript per million
TRIM	Tripartite motif containing
TTP	Tristetraprolin

VEGF	Vascular endothelial growth factor
UTR	Untranslated region
ZEB1	Zinc finger E-box binding homeobox 1

BIBLIOGRAPHY

1. Gjorevski, N. and C.M. Nelson, *Integrated morphodynamic signalling of the mammary gland.* Nat Rev Mol Cell Biol, 2011. **12**(9): p. 581-93.
2. Chatterjee, S.J. and L. McCaffrey, *Emerging role of cell polarity proteins in breast cancer progression and metastasis.* Breast Cancer (Dove Med Press), 2014. **6**: p. 15-27.
3. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2018.* CA Cancer J Clin, 2018. **68**(1): p. 7-30.
4. Guise, T.A., *Breast cancer bone metastases: it's all about the neighborhood.* Cell, 2013. **154**(5): p. 957-959.
5. Minn, A.J., et al., *Genes that mediate breast cancer metastasis to lung.* Nature, 2005. **436**(7050): p. 518-24.
6. DeSantis, C.E., et al., *Breast cancer statistics, 2017, racial disparity in mortality by state.* CA Cancer J Clin, 2017. **67**(6): p. 439-448.
7. Sun, Y.S., et al., *Risk Factors and Preventions of Breast Cancer.* Int J Biol Sci, 2017. **13**(11): p. 1387-1397.
8. DeSantis, C., et al., *Breast cancer statistics, 2013.* CA Cancer J Clin, 2014. **64**(1): p. 52-62.
9. Shi, J., et al., *Disrupting the interaction of BRD4 with diacetylated Twist suppresses tumorigenesis in basal-like breast cancer.* Cancer Cell, 2014. **25**(2): p. 210-25.
10. Gao, J.J. and S.M. Swain, *Luminal A Breast Cancer and Molecular Assays: A Review.* Oncologist, 2018. **23**(5): p. 556-565.
11. Rakha, E.A. and A.R. Green, *Molecular classification of breast cancer: what the pathologist needs to know.* Pathology, 2017. **49**(2): p. 111-119.
12. Shi, J., J. Cao, and B.P. Zhou, *Twist-BRD4 complex: potential drug target for basal-like breast cancer.* Curr Pharm Des, 2015. **21**(10): p. 1256-61.
13. Jamdade, V.S., et al., *Therapeutic targets of triple-negative breast cancer: a review.* Br J Pharmacol, 2015. **172**(17): p. 4228-37.
14. Leidy, J., A. Khan, and D. Kandil, *Basal-like breast cancer: update on clinicopathologic, immunohistochemical, and molecular features.* Arch Pathol Lab Med, 2014. **138**(1): p. 37-43.
15. Toft, D.J. and V.L. Cryns, *Minireview: Basal-like breast cancer: from molecular profiles to targeted therapies.* Mol Endocrinol, 2011. **25**(2): p. 199-211.

16. Tassone, P., et al., *BRCA1 expression modulates chemosensitivity of BRCA1-defective HCC1937 human breast cancer cells.* Br J Cancer, 2003. **88**(8): p. 1285-91.
17. Botti, G., et al., *Programmed Death Ligand 1 (PD-L1) Tumor Expression Is Associated with a Better Prognosis and Diabetic Disease in Triple Negative Breast Cancer Patients.* Int J Mol Sci, 2017. **18**(2).
18. Marra, A., G. Viale, and G. Curigliano, *Recent advances in triple negative breast cancer: the immunotherapy era.* BMC Med, 2019. **17**(1): p. 90.
19. Zhou, Y., E.B. Rucker, 3rd, and B.P. Zhou, *Autophagy regulation in the development and treatment of breast cancer.* Acta Biochim Biophys Sin (Shanghai), 2016. **48**(1): p. 60-74.
20. Luzzi, K.J., et al., *Multistep nature of metastatic inefficiency: dormancy of solitary cells after successful extravasation and limited survival of early micrometastases.* Am J Pathol, 1998. **153**(3): p. 865-73.
21. Kimbung, S., N. Loman, and I. Hedenfalk, *Clinical and molecular complexity of breast cancer metastases.* Semin Cancer Biol, 2015. **35**: p. 85-95.
22. Wan, L., K. Pantel, and Y. Kang, *Tumor metastasis: moving new biological insights into the clinic.* Nat Med, 2013. **19**(11): p. 1450-64.
23. Xu, J., S. Lamouille, and R. Derynck, *TGF-beta-induced epithelial to mesenchymal transition.* Cell Res, 2009. **19**(2): p. 156-72.
24. Wang, Y. and B.P. Zhou, *Epithelial-mesenchymal transition in breast cancer progression and metastasis.* Chin J Cancer, 2011. **30**(9): p. 603-11.
25. Lindsey, S. and S.A. Langhans, *Crosstalk of Oncogenic Signaling Pathways during Epithelial-Mesenchymal Transition.* Front Oncol, 2014. **4**: p. 358.
26. Polyak, K. and R.A. Weinberg, *Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits.* Nat Rev Cancer, 2009. **9**(4): p. 265-73.
27. Nassar, D. and C. Blanpain, *Cancer Stem Cells: Basic Concepts and Therapeutic Implications.* Annu Rev Pathol, 2016. **11**: p. 47-76.
28. Soon, P.S., et al., *Breast cancer-associated fibroblasts induce epithelial-to-mesenchymal transition in breast cancer cells.* Endocr Relat Cancer, 2013. **20**(1): p. 1-12.
29. Dongre, A. and R.A. Weinberg, *New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer.* Nat Rev Mol Cell Biol, 2019. **20**(2): p. 69-84.

30. Kalluri, R. and R.A. Weinberg, *The basics of epithelial-mesenchymal transition.* J Clin Invest, 2009. **119**(6): p. 1420-8.
31. Lintz, M., A. Munoz, and C.A. Reinhart-King, *The Mechanics of Single Cell and Collective Migration of Tumor Cells.* J Biomech Eng, 2017. **139**(2).
32. Friedl, P., *Prespecification and plasticity: shifting mechanisms of cell migration.* Curr Opin Cell Biol, 2004. **16**(1): p. 14-23.
33. Hecht, I., et al., *Tumor invasion optimization by mesenchymal-amoeboid heterogeneity.* Sci Rep, 2015. **5**: p. 10622.
34. Wolf, K., et al., *Physical limits of cell migration: control by ECM space and nuclear deformation and tuning by proteolysis and traction force.* J Cell Biol, 2013. **201**(7): p. 1069-84.
35. Mak, M., et al., *Single-Cell Migration in Complex Microenvironments: Mechanics and Signaling Dynamics.* J Biomech Eng, 2016. **138**(2): p. 021004.
36. Aiello, N.M., et al., *EMT Subtype Influences Epithelial Plasticity and Mode of Cell Migration.* Dev Cell, 2018. **45**(6): p. 681-695 e4.
37. Nieto, M.A., et al., *Emt: 2016.* Cell, 2016. **166**(1): p. 21-45.
38. Giampieri, S., et al., *Localized and reversible TGFbeta signalling switches breast cancer cells from cohesive to single cell motility.* Nat Cell Biol, 2009. **11**(11): p. 1287-96.
39. Aceto, N., et al., *Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis.* Cell, 2014. **158**(5): p. 1110-1122.
40. Cheung, K.J. and A.J. Ewald, *A collective route to metastasis: Seeding by tumor cell clusters.* Science, 2016. **352**(6282): p. 167-9.
41. Maddipati, R. and B.Z. Stanger, *Pancreatic Cancer Metastases Harbor Evidence of Polyclonality.* Cancer Discov, 2015. **5**(10): p. 1086-97.
42. Reichert, M., et al., *Regulation of Epithelial Plasticity Determines Metastatic Organotropism in Pancreatic Cancer.* Dev Cell, 2018. **45**(6): p. 696-711 e8.
43. Scheel, C., et al., *Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast.* Cell, 2011. **145**(6): p. 926-40.
44. Su, J., et al., *MicroRNA-200a suppresses the Wnt/beta-catenin signaling pathway by interacting with beta-catenin.* Int J Oncol, 2012. **40**(4): p. 1162-70.
45. Chaffer, C.L., et al., *EMT, cell plasticity and metastasis.* Cancer Metastasis Rev, 2016. **35**(4): p. 645-654.
46. Nagathihalli, N.S. and N.B. Merchant, *Src-mediated regulation of E-cadherin and EMT in pancreatic cancer.* Front Biosci (Landmark Ed), 2012. **17**: p. 2059-69.

47. Batlle, E., et al., *The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells.* Nat Cell Biol, 2000. **2**(2): p. 84-9.
48. Lamouille, S., J. Xu, and R. Derynck, *Molecular mechanisms of epithelial-mesenchymal transition.* Nat Rev Mol Cell Biol, 2014. **15**(3): p. 178-96.
49. Barrallo-Gimeno, A. and M.A. Nieto, *The Snail genes as inducers of cell movement and survival: implications in development and cancer.* Development, 2005. **132**(14): p. 3151-61.
50. Yang, J., et al., *Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis.* Cell, 2004. **117**(7): p. 927-39.
51. Aparicio, L.A., et al., *Posttranscriptional regulation by RNA-binding proteins during epithelial-to-mesenchymal transition.* Cell Mol Life Sci, 2013. **70**(23): p. 4463-77.
52. de Moor, C.H., H. Meijer, and S. Lissenden, *Mechanisms of translational control by the 3' UTR in development and differentiation.* Semin Cell Dev Biol, 2005. **16**(1): p. 49-58.
53. Sutherland, J.M., et al., *RNA binding proteins in spermatogenesis: an in depth focus on the Musashi family.* Asian J Androl, 2015. **17**(4): p. 529-36.
54. Massague, J., *TGFbeta in Cancer.* Cell, 2008. **134**(2): p. 215-30.
55. Imamura, T., A. Hikita, and Y. Inoue, *The roles of TGF-beta signaling in carcinogenesis and breast cancer metastasis.* Breast Cancer, 2012. **19**(2): p. 118-24.
56. Principe, D.R., et al., *TGF-beta: duality of function between tumor prevention and carcinogenesis.* J Natl Cancer Inst, 2014. **106**(2): p. djt369.
57. He, W., et al., *Hematopoiesis controlled by distinct TIF1gamma and Smad4 branches of the TGFbeta pathway.* Cell, 2006. **125**(5): p. 929-41.
58. Zhang, L., et al., *RNF12 controls embryonic stem cell fate and morphogenesis in zebrafish embryos by targeting Smad7 for degradation.* Mol Cell, 2012. **46**(5): p. 650-61.
59. Koinuma, D., et al., *Arkadia amplifies TGF-beta superfamily signalling through degradation of Smad7.* EMBO J, 2003. **22**(24): p. 6458-70.
60. Derynck, R. and Y.E. Zhang, *Smad-dependent and Smad-independent pathways in TGF-beta family signalling.* Nature, 2003. **425**(6958): p. 577-84.
61. Zhang, Y.E., *Non-Smad Signaling Pathways of the TGF-beta Family.* Cold Spring Harb Perspect Biol, 2017. **9**(2).

62. Yi, J.Y., I. Shin, and C.L. Arteaga, *Type I transforming growth factor beta receptor binds to and activates phosphatidylinositol 3-kinase.* J Biol Chem, 2005. **280**(11): p. 10870-6.
63. Sorrentino, A., et al., *The type I TGF-beta receptor engages TRAF6 to activate TAK1 in a receptor kinase-independent manner.* Nat Cell Biol, 2008. **10**(10): p. 1199-207.
64. Bhowmick, N.A., et al., *Transforming growth factor-beta1 mediates epithelial to mesenchymal transdifferentiation through a RhoA-dependent mechanism.* Mol Biol Cell, 2001. **12**(1): p. 27-36.
65. Edlund, S., et al., *Transforming growth factor-beta-induced mobilization of actin cytoskeleton requires signaling by small GTPases Cdc42 and RhoA.* Mol Biol Cell, 2002. **13**(3): p. 902-14.
66. Xie, F., et al., *TGF-beta signaling in cancer metastasis.* Acta Biochim Biophys Sin (Shanghai), 2018. **50**(1): p. 121-132.
67. Gordon, K.J. and G.C. Blobe, *Role of transforming growth factor-beta superfamily signaling pathways in human disease.* Biochim Biophys Acta, 2008. **1782**(4): p. 197-228.
68. Derynck, R. and R.J. Akhurst, *Differentiation plasticity regulated by TGF-beta family proteins in development and disease.* Nat Cell Biol, 2007. **9**(9): p. 1000-4.
69. Wahl, S.M., J. Wen, and N. Moutsopoulos, *TGF-beta: a mobile purveyor of immune privilege.* Immunol Rev, 2006. **213**: p. 213-27.
70. Horbelt, D., A. Denkis, and P. Knaus, *A portrait of Transforming Growth Factor beta superfamily signalling: Background matters.* Int J Biochem Cell Biol, 2012. **44**(3): p. 469-74.
71. Wang, Y. and B.P. Zhou, *Epithelial-mesenchymal Transition---A Hallmark of Breast Cancer Metastasis.* Cancer Hallm, 2013. **1**(1): p. 38-49.
72. Klaus, A. and W. Birchmeier, *Wnt signalling and its impact on development and cancer.* Nat Rev Cancer, 2008. **8**(5): p. 387-98.
73. Clevers, H., *Wnt/beta-catenin signaling in development and disease.* Cell, 2006. **127**(3): p. 469-80.
74. Gauger, K.J., et al., *SFRP1 reduction results in an increased sensitivity to TGF-beta signaling.* BMC Cancer, 2011. **11**: p. 59.
75. Tang, Y., et al., *Smad7 stabilizes beta-catenin binding to E-cadherin complex and promotes cell-cell adhesion.* J Biol Chem, 2008. **283**(35): p. 23956-63.
76. Hoover, L.L. and S.W. Kubalak, *Holding their own: the noncanonical roles of Smad proteins.* Sci Signal, 2008. **1**(46): p. pe48.

77. Dissanayake, S.K., et al., *The Wnt5A/protein kinase C pathway mediates motility in melanoma cells via the inhibition of metastasis suppressors and initiation of an epithelial to mesenchymal transition.* J Biol Chem, 2007. **282**(23): p. 17259-71.
78. Kopan, R., *Notch: a membrane-bound transcription factor.* J Cell Sci, 2002. **115**(Pt 6): p. 1095-7.
79. Mollen, E.W.J., et al., *Moving Breast Cancer Therapy up a Notch.* Front Oncol, 2018. **8**: p. 518.
80. Kopan, R. and M.X. Ilagan, *The canonical Notch signaling pathway: unfolding the activation mechanism.* Cell, 2009. **137**(2): p. 216-33.
81. Wang, K., et al., *PEST domain mutations in Notch receptors comprise an oncogenic driver segment in triple-negative breast cancer sensitive to a gamma-secretase inhibitor.* Clin Cancer Res, 2015. **21**(6): p. 1487-96.
82. Dontu, G., et al., *Role of Notch signaling in cell-fate determination of human mammary stem/progenitor cells.* Breast Cancer Res, 2004. **6**(6): p. R605-15.
83. Bouras, T., et al., *Notch signaling regulates mammary stem cell function and luminal cell-fate commitment.* Cell Stem Cell, 2008. **3**(4): p. 429-41.
84. Xie, M., et al., *Activation of Notch-1 enhances epithelial-mesenchymal transition in gefitinib-acquired resistant lung cancer cells.* J Cell Biochem, 2012. **113**(5): p. 1501-13.
85. Zhang, L., et al., *Activation of Notch pathway is linked with epithelial-mesenchymal transition in prostate cancer cells.* Cell Cycle, 2017. **16**(10): p. 999-1007.
86. Gao, X.J., et al., *Nobiletin inhibited hypoxia-induced epithelial-mesenchymal transition of lung cancer cells by inactivating of Notch-1 signaling and switching on miR-200b.* Pharmazie, 2015. **70**(4): p. 256-62.
87. Bao, B., et al., *Notch-1 induces epithelial-mesenchymal transition consistent with cancer stem cell phenotype in pancreatic cancer cells.* Cancer Lett, 2011. **307**(1): p. 26-36.
88. Timmerman, L.A., et al., *Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation.* Genes Dev, 2004. **18**(1): p. 99-115.
89. Zhao, Z.L., et al., *Notch signaling induces epithelial-mesenchymal transition to promote invasion and metastasis in adenoid cystic carcinoma.* Am J Transl Res, 2015. **7**(1): p. 162-74.
90. Kim, R.K., et al., *Radiation driven epithelial-mesenchymal transition is mediated by Notch signaling in breast cancer.* Oncotarget, 2016. **7**(33): p. 53430-53442.

91. Ito, T., et al., *Small cell lung cancer, an epithelial to mesenchymal transition (EMT)-like cancer: significance of inactive Notch signaling and expression of achaete-scute complex homologue 1.* Hum Cell, 2017. **30**(1): p. 1-10.
92. Saad, S., et al., *Notch mediated epithelial to mesenchymal transformation is associated with increased expression of the Snail transcription factor.* Int J Biochem Cell Biol, 2010. **42**(7): p. 1115-22.
93. Zhang, J., et al., *NUMB negatively regulates the epithelial-mesenchymal transition of triple-negative breast cancer by antagonizing Notch signaling.* Oncotarget, 2016. **7**(38): p. 61036-61053.
94. Sczaniecka, M., et al., *MDM2 protein-mediated ubiquitination of numb protein: identification of a second physiological substrate of MDM2 that employs a dual-site docking mechanism.* J Biol Chem, 2012. **287**(17): p. 14052-68.
95. Di Domenico, M. and A. Giordano, *Signal transduction growth factors: the effective governance of transcription and cellular adhesion in cancer invasion.* Oncotarget, 2017. **8**(22): p. 36869-36884.
96. Colomiere, M., et al., *Cross talk of signals between EGFR and IL-6R through JAK2/STAT3 mediate epithelial-mesenchymal transition in ovarian carcinomas.* Br J Cancer, 2009. **100**(1): p. 134-44.
97. Kim, J., et al., *EGF induces epithelial-mesenchymal transition through phospho-Smad2/3-Snail signaling pathway in breast cancer cells.* Oncotarget, 2016. **7**(51): p. 85021-85032.
98. Shirakihara, T., et al., *TGF-beta regulates isoform switching of FGF receptors and epithelial-mesenchymal transition.* EMBO J, 2011. **30**(4): p. 783-95.
99. Lei, H. and C.X. Deng, *Fibroblast Growth Factor Receptor 2 Signaling in Breast Cancer.* Int J Biol Sci, 2017. **13**(9): p. 1163-1171.
100. Ogunwobi, O.O., et al., *Epigenetic upregulation of HGF and c-Met drives metastasis in hepatocellular carcinoma.* PLoS One, 2013. **8**(5): p. e63765.
101. Tam, W.L., et al., *Protein kinase C alpha is a central signaling node and therapeutic target for breast cancer stem cells.* Cancer Cell, 2013. **24**(3): p. 347-64.
102. Audic, Y. and R.S. Hartley, *Post-transcriptional regulation in cancer.* Biol Cell, 2004. **96**(7): p. 479-98.
103. Kondo, Y., et al., *Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5' splice site recognition.* Elife, 2015. **4**.

104. Lorkovic, Z.J., et al., *Evolutionary conservation of minor U12-type spliceosome between plants and humans.* RNA, 2005. **11**(7): p. 1095-107.
105. Wang, Y., et al., *Mechanism of alternative splicing and its regulation.* Biomed Rep, 2015. **3**(2): p. 152-158.
106. Busch, A. and K.J. Hertel, *Evolution of SR protein and hnRNP splicing regulatory factors.* Wiley Interdiscip Rev RNA, 2012. **3**(1): p. 1-12.
107. Pagliarini, V., C. Naro, and C. Sette, *Splicing Regulation: A Molecular Device to Enhance Cancer Cell Adaptation.* Biomed Res Int, 2015. **2015**: p. 543067.
108. Liu, Y., et al., *Impact of Alternative Splicing on the Human Proteome.* Cell Rep, 2017. **20**(5): p. 1229-1241.
109. Lopez de Silanes, I., M.P. Quesada, and M. Esteller, *Aberrant regulation of messenger RNA 3'-untranslated region in human cancer.* Cell Oncol, 2007. **29**(1): p. 1-17.
110. Biamonti, G., et al., *Making alternative splicing decisions during epithelial-to-mesenchymal transition (EMT).* Cell Mol Life Sci, 2012. **69**(15): p. 2515-26.
111. Semmler, L., C. Reiter-Brennan, and A. Klein, *BRCA1 and Breast Cancer: a Review of the Underlying Mechanisms Resulting in the Tissue-Specific Tumorigenesis in Mutation Carriers.* J Breast Cancer, 2019. **22**(1): p. 1-14.
112. Silipo, M., H. Gautrey, and A. Tyson-Capper, *Deregulation of splicing factors and breast cancer development.* J Mol Cell Biol, 2015. **7**(5): p. 388-401.
113. Orban, T.I. and E. Olah, *Emerging roles of BRCA1 alternative splicing.* Mol Pathol, 2003. **56**(4): p. 191-7.
114. Cartegni, L., S.L. Chew, and A.R. Krainer, *Listening to silence and understanding nonsense: exonic mutations that affect splicing.* Nat Rev Genet, 2002. **3**(4): p. 285-98.
115. Moore, M.J., *From birth to death: the complex lives of eukaryotic mRNAs.* Science, 2005. **309**(5740): p. 1514-8.
116. Griseri, P. and G. Pages, *Regulation of the mRNA half-life in breast cancer.* World J Clin Oncol, 2014. **5**(3): p. 323-34.
117. Andreassi, C. and A. Riccio, *To localize or not to localize: mRNA fate is in 3'UTR ends.* Trends Cell Biol, 2009. **19**(9): p. 465-74.
118. Mayr, C. and D.P. Bartel, *Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells.* Cell, 2009. **138**(4): p. 673-84.
119. Di Giammartino, D.C., K. Nishida, and J.L. Manley, *Mechanisms and consequences of alternative polyadenylation.* Mol Cell, 2011. **43**(6): p. 853-66.

120. Ghigna, C., et al., *Posttranscriptional Regulation and RNA Binding Proteins in Cancer Biology.* Biomed Res Int, 2015. **2015**: p. 897821.
121. Anantharaman, V., E.V. Koonin, and L. Aravind, *Comparative genomics and evolution of proteins involved in RNA metabolism.* Nucleic Acids Res, 2002. **30**(7): p. 1427-64.
122. Wang, W., et al., *HuR regulates cyclin A and cyclin B1 mRNA stability during cell proliferation.* EMBO J, 2000. **19**(10): p. 2340-50.
123. Dong, R., et al., *Stabilization of Snail by HuR in the process of hydrogen peroxide induced cell migration.* Biochem Biophys Res Commun, 2007. **356**(1): p. 318-21.
124. Maltzman, W. and L. Czyzyk, *UV irradiation stimulates levels of p53 cellular tumor antigen in nontransformed mouse cells.* Mol Cell Biol, 1984. **4**(9): p. 1689-94.
125. Gratacos, F.M. and G. Brewer, *The role of AUF1 in regulated mRNA decay.* Wiley Interdiscip Rev RNA, 2010. **1**(3): p. 457-73.
126. Ciais, D., N. Cherradi, and J.J. Feige, *Multiple functions of tristetraprolin/TIS11 RNA-binding proteins in the regulation of mRNA biogenesis and degradation.* Cell Mol Life Sci, 2013. **70**(12): p. 2031-44.
127. Meisner, N.C., et al., *mRNA openers and closers: modulating AU-rich element-controlled mRNA stability by a molecular switch in mRNA secondary structure.* Chembiochem, 2004. **5**(10): p. 1432-47.
128. Chou, C.F., et al., *Tethering KSRP, a decay-promoting AU-rich element-binding protein, to mRNAs elicits mRNA decay.* Mol Cell Biol, 2006. **26**(10): p. 3695-706.
129. Penkov, D., et al., *Cloning of a human gene closely related to the genes coding for the c-myc single-strand binding proteins.* Gene, 2000. **243**(1-2): p. 27-36.
130. Wu, Y., et al., *Down regulation of RNA binding motif, single-stranded interacting protein 3, along with up regulation of nuclear HIF1A correlates with poor prognosis in patients with gastric cancer.* Oncotarget, 2017. **8**(1): p. 1262-1277.
131. Chen, J., et al., *RBMS3 at 3p24 inhibits nasopharyngeal carcinoma development via inhibiting cell proliferation, angiogenesis, and inducing apoptosis.* PLoS One, 2012. **7**(9): p. e44636.
132. Zhang, T., et al., *Low expression of RBMS3 and SFRP1 are associated with poor prognosis in patients with gastric cancer.* Am J Cancer Res, 2016. **6**(11): p. 2679-2689.
133. Fritz, D. and B. Stefanovic, *RNA-binding protein RBMS3 is expressed in activated hepatic stellate cells and liver fibrosis and increases expression of transcription factor Prx1.* J Mol Biol, 2007. **371**(3): p. 585-95.

134. Wu, G., et al., *Loss of RBMS3 Confers Platinum-resistance in Epithelial Ovarian Cancer via Activation of miR-126-5p/beta-catenin/CBP signaling.* Clin Cancer Res, 2018.
135. Lu, C.K., et al., *Rbms3, an RNA-binding protein, mediates the expression of Ptf1a by binding to its 3'UTR during mouse pancreas development.* DNA Cell Biol, 2012. **31**(7): p. 1245-51.
136. Jayasena, C.S. and M.E. Bronner, *Rbms3 functions in craniofacial development by posttranscriptionally modulating TGF-beta signaling.* J Cell Biol, 2012. **199**(3): p. 453-66.
137. Song, J., *EMT or apoptosis: a decision for TGF-beta.* Cell Res, 2007. **17**(4): p. 289-90.
138. Grimshaw, M.J., et al., *Mammosphere culture of metastatic breast cancer cells enriches for tumorigenic breast cancer cells.* Breast Cancer Res, 2008. **10**(3): p. R52.
139. Wang, J., et al., *TSPAN31 is a critical regulator on transduction of survival and apoptotic signals in hepatocellular carcinoma cells.* FEBS Lett, 2017. **591**(18): p. 2905-2918.
140. Dong, C., et al., *Loss of FBP1 by Snail-mediated repression provides metabolic advantages in basal-like breast cancer.* Cancer Cell, 2013. **23**(3): p. 316-31.
141. Hong, Y., et al., *PPARgamma mediates the effects of WIN55,212-2, an synthetic cannabinoid, on the proliferation and apoptosis of the BEL-7402 hepatocarcinoma cells.* Mol Biol Rep, 2013. **40**(11): p. 6287-93.
142. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.* Nat Genet, 2000. **25**(1): p. 25-9.
143. The Gene Ontology, C., *The Gene Ontology Resource: 20 years and still GOing strong.* Nucleic Acids Res, 2019. **47**(D1): p. D330-D338.
144. Zhou, Y., et al., *Metascape provides a biologist-oriented resource for the analysis of systems-level datasets.* Nat Commun, 2019. **10**(1): p. 1523.
145. Gyorffy, B., et al., *An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients.* Breast Cancer Res Treat, 2010. **123**(3): p. 725-31.
146. Mani, S.A., et al., *The epithelial-mesenchymal transition generates cells with properties of stem cells.* Cell, 2008. **133**(4): p. 704-15.
147. Singh, A. and J. Settleman, *EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer.* Oncogene, 2010. **29**(34): p. 4741-51.

148. Morel, A.P., et al., *Generation of breast cancer stem cells through epithelial-mesenchymal transition.* PLoS One, 2008. **3**(8): p. e2888.
149. Chang, J.C., *Cancer stem cells: Role in tumor growth, recurrence, metastasis, and treatment resistance.* Medicine (Baltimore), 2016. **95**(1 Suppl 1): p. S20-5.
150. Massague, J., *A very private TGF-beta receptor embrace.* Mol Cell, 2008. **29**(2): p. 149-50.
151. Tang, X., et al., *SIRT7 antagonizes TGF-beta signaling and inhibits breast cancer metastasis.* Nat Commun, 2017. **8**(1): p. 318.
152. Bebee, T.W., et al., *The splicing regulators Esrp1 and Esrp2 direct an epithelial splicing program essential for mammalian development.* Elife, 2015. **4**.
153. Ishii, H., et al., *Epithelial splicing regulatory proteins 1 (ESRP1) and 2 (ESRP2) suppress cancer cell motility via different mechanisms.* J Biol Chem, 2014. **289**(40): p. 27386-99.
154. Warzecha, C.C., et al., *ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing.* Mol Cell, 2009. **33**(5): p. 591-601.
155. Warzecha, C.C., et al., *The epithelial splicing factors ESRP1 and ESRP2 positively and negatively regulate diverse types of alternative splicing events.* RNA Biol, 2009. **6**(5): p. 546-62.
156. Furumatsu, T., et al., *Smad3 induces chondrogenesis through the activation of SOX9 via CREB-binding protein/p300 recruitment.* J Biol Chem, 2005. **280**(9): p. 8343-50.
157. Chiechi, A., et al., *Role of TGF-beta in breast cancer bone metastases.* Adv Biosci Biotechnol, 2013. **4**(10C): p. 15-30.
158. Ikushima, H. and K. Miyazono, *TGFbeta signalling: a complex web in cancer progression.* Nat Rev Cancer, 2010. **10**(6): p. 415-24.
159. Kim, B., et al., *A CTGF-RUNX2-RANKL Axis in Breast and Prostate Cancer cells Promotes Tumor Progression in Bone.* J Bone Miner Res, 2019.

VITA

Education

2015-Present Ph.D. in Molecular and Cellular Biochemistry, University of Kentucky

2012–2014 M.S. in Molecular, Cellular and Developmental Biology, University of Kentucky

2008-2012 B.S. in Biotechnology, South Medical University, China

Professional Experience

2012-2014 Teaching Assistant in Intro to Microbiology, Principles of Genetics and Principles of Biology II, University of Kentucky

Academic Award

2019 Selected as the Representative for the University of Kentucky to Apply for the National Cancer Institute-Predoctoral to Postdoctoral Transition Fellowship (F99/K00)

2015-Present Graduate Assistantship, University of Kentucky

2012-2014 Teaching Assistantship, University of Kentucky

2011 South Medical University Undergraduate Scholarship Award

2011 First Place in Comprehensive Quality Assessment, School of Biotechnology, South Medical University, China

2008-2009 Merit Student, South Medical University, China

2008 Top Ten Excellent Female Student Award, South Medical University, China

Publication

Zhou YT, Rucker E, Zhou BP. Autophagy regulation in the development and treatment of breast cancer. *Acta Biochim Biophys Sin(Shanghai)*, 2016 Jan 48(1):60-74

Hong YH, **Zhou YT**, et al. PPARγ mediates WIN55,212-2 induced apoptosis of BEL-7402 hepatocarcinoma cells. *Molecular Biology Report*, 2013.

Liu YJ, Li H, **Zhou YT**, et al. Procaryotic Expression and Multi-Clonal Preparation of PfRPA2. *Journal of Tropical Medicine,* 2012.12(6)

Xiao SH, **Zhou YT**, et al. Effect of Activation of Cannabinoid Receptor by THC on Proliferation and Apoptosis of Hepatoma Bel-7402 Cells. *Lishizhen Medicine and Materia Medica Research,* 2012. 23(5)

Zhu XQ, Hu JX, **Zhou YT,** et al. Effect of the cannabinoid receptor activation by THC on proliferation and apoptosis of cancer A549 cells. *Acta Med Univ Sci Technol Huazhong*, 2011;40(5):354-358

Wang Y, **Zhou YT**, et al. The effect of the activation of cannabinoid receptor WIN-55 212-2 on the proliferation and apoptosis of hepatoma HepG2 cells. *Chinese J of Cell and Molecular Immunol*, 2010;26(4): 344-348

Zhu XQ, **Zhou YT**, et al. The effect of the activation of cannabinoid receptor THC on the proliferation and apoptosis of hepatoma HepG2 cells. *Acta Med Univ Sci Technol Huazhong*, 2010;39(3): 376-380

Oral Presentation

2018 "RBMS3 Promotes Basal-like Breast Cancer by Regulating TGF-β Signaling"

Data Club Presentation

Department of Molecular and Cellular Biochemistry, University of Kentucky

2018 "The Role of RNA Binding Protein RBMS3 in Basal-like Breast Cancer"

Departmental Seminar Presentation

Department of Molecular and Cellular Biochemistry, University of Kentucky

2017 "The Role of Glucocorticoid Receptor-Mediated Hippo Signaling in Breast Cancer Cell Stemness and Chemoresistance"

Departmental Seminar Presentation

Department of Molecular and Cellular Biochemistry, University of Kentucky

2016 "Impaired SIRT1 Activity and Huntington's Disease"

Departmental Seminar Presentation

Department of Molecular and Cellular Biochemistry, University of Kentucky

2015 "lncRNA *BCAR4* and Breast Cancer Metastasis"

Departmental Seminar Presentation

Department of Molecular and Cellular Biochemistry, University of Kentucky

2013 "PPARγ Mediates WIN55,212-2, a Synthetic Cannabinoid, on the Proliferation and Apoptosis of Liver Cancer Cells"

The Sixth Annual World Cancer Congress, Xi'An, Shanxi, China